N° ORDRE : 40381

École doctorale régionale Sciences Pour l'Ingénieur Lille Nord–de–France

Université Lille 1 Sciences et Technologies

MODÉLISATION, ANALYSE MATHÉMATIQUE ET SIMULATION NUMÉRIQUE DE PROBLÈMES ISSUS DE LA BIOLOGIE

THÈSE

présentée et soutenue publiquement le 7 décembre 2010

pour l'obtention du

Doctorat de l'université Lille 1

(spécialité mathématiques)

par

ANNE DEVYS

Composition du jury

Président :	Jean–Frédéric GERBEAU	INRIA Paris–Rocquencourt
Rapporteurs :	Thierry COLIN	Université de Bordeaux 1
	Stéphane DESCOMBES	Université de Nice
Examinateurs :	Claire CHAINAIS	Université de Lille 1
	Thierry GOUDON	INRIA Lille–Nord Europe *(directeur)*
	Pauline LAFITTE	Université de Lille 1 *(directrice)*

Laboratoire Paul Painlevé — **Equipe SIMPAF (INRIA Lille Nord–Europe)**

Mis en page avec la classe thloria.

Merci !

Tout d'abord merci à Pauline et Thierry pour m'avoir ouvert les portes du monde de la recherche et encouragée à le parcourir en long, en large et en travers.

Je remercie mes deux rapporteurs MONSIEUR COLIN et MONSIEUR DESCOMBES pour l'intérêt qu'ils ont apporté à mon manuscrit, ainsi que MADAME CHAINAIS, et MONSIEUR GERBEAU d'avoir accepté de faire partie du jury.

Je suis également reconnaissante envers le Laboratoire Paul Painlevé et le centre INRIA Lille–Nord–Europe, qui ont réglé tous les aspects matériels pendant ces trois ans. Merci en particulier à Sandrine, notre assistante, Soledad et Valérie, que j'ai dérangées tant de fois pour mes ordres de mission si nombreux en plus des invitations au séminaire des doctorants. Merci aussi aux membres des équipes Analyse numérique et Équations aux dérivées partielles et SIMPAF pour leur accueil.

Cette thèse est le fruit d'un jeu à l'allemande : offensif certe, mais surtout collectif. Nous étions un peu plus de onze sur le terrain (oui, c'est de la triche...), mais ce fut un grand plaisir de travailler avec des amis dans la joie et surtout la bonne humeur.

Ces dernières années j'ai eu l'occasion de traîner mes guêtres un peu partout, trouvant à chaque fois un nouveau 'chef', une nouvelle famille. Merci à tous mes chefs de France et d'ailleurs, A. Debussche (le premier), Céline, pour laquelle le terme de mathématicienne prend tout son sens, Bertrand, 'Big Chef', Sébastien, mon meilleur soutien, Vincent et José–Antonio, que de bons souvenirs des moments passés à Barcelone (Vincent est le meilleur guide touristique de la ville que je connaisse).

Merci à tous mes collaborateurs : Driss, mon premier grand collaborateur, patient, prévenant et adorable, parti chasser l'ours polaire au Canada, Bérénice, toujours souriante et encourageante, avec qui les pauses BN deviennent des moments inoubliables et Ralf, le 'Monsieur Bio' de notre équipe Lilloise.

Je remercie Grégory et Virginie sources rennaises de bonne humeur. Je me souviendrai aussi longtemps d'Erwann expliquant à cette pigiste de Ouest–France que les mathématiques c'était toute une philosophie de vie.

Des gros bisous pour tous les membres de la famille Cemracs, en particulier au plus illustre d'entre eux, Mister Cemracs, qui en plus de toutes les qualités qui lui ont valu son titre a été une oreille attentive. Merci beaucoup Frédéric. Merci à la fine équipe : Laurent, indissociable de son téléphone (Tu reprendras bien un verre de lait ?), Dario, qui se souviendra sans doute longtemps de ses leçons de français marseillaises. Merci à Vanessa (Madame Picon), Hélène (Miss Cemracs), Jean-Baptiste, Philippe, Stéphane, Ulrich, Matteo, Cristobal, Paul, un habitué comme moi des urgences de Ste Marguerite, Amélie, qui tient un 4 étoiles au centre de Lyon, Nicolas S., soirée inoubliable chez la Clodette, Ayman, qui n'hésita pas une seconde à me céder son badge dans l'espoir de me sortir d'une situation inconfortable, Alexis B., Pierre G., Séverine, Frédérique, Federico et bien entendu Jean-Marie, Carlos, Laurence et Nordin.

Je remercie encore Adrien, Petit Ludo et Grand Ludo, Benoît P., Gros Ben et Lima, toute une lignée de tré-

soriers, Elodie et Mathieu (les amoureux), Maud, ma binôme préférée, Raphaël R., le petit Mickaël, la grande Marie, Villeparisis a été mon meilleur refuge, la petite Marie, qui m'expliqua les MAPK autrefois, Jon, Aurélien, Mathieu L.R., Arnault, Camille P., le très accueillant 2D24, Vincent B., qui soigna ma peur du probabiliste, Benoît, Adeline et Aude pour leur accueil à Orsay, Jesús et José–Alfredo, le garage Perthame. Merci beaucoup à mes guides, Alexis G. et Benoît G. Merci à Alexis pour les sorties dans les bars les plus *Select* de Lille. Et Benoît, merci de m'avoir donné l'occasion de te reprendre dans ton explication du hors–jeu. Merci encore à Shuyan, Raphaël L-R., Vincent V., Nicolas V., le grand frère, Martin, adoubé récemment 'chef du bureau des rigolos', Léon (le cochon), Benjamin, Chang, Éric, le chef des doctorants, Oscar, mon mari. Merci particulier à Bénédicte, compagne de misère, encore invaincue au *Shabadabada* et fan des moustachus. Devant toutes mes bêtises, tu as été la sagesse qui a fait que je sois encore dans le presque droit chemin. Merci à Mathias,

Toi qui dessine au fond de ton ennui les notes d'une mélodie
Une musique sans accord majeur c'est une piste sans danseur
Mais si tu ranges ces moments–là dans un placard tes idées noires
Les notes pourront se danser et nous reviendrons les chanter...

J'ai parié que nous allions être bientôt conviés à l'enterrement de Germain et Germaine, j'espère que tu me feras mentir (encore une fois). Merci au très piquant Guillaume, roi de la gaine électrique pour qui j'ai été jusqu'à me mettre à genou dans la poussière. Merci Herrn Doktor Antoine Gloria, Jean–François qui sait tout, Ingrid et sa "toujours partante attitude", Frédéric B., Juliette et tout le club des Voedtsettes, Pierre P., tuteur pédagogique parfait avec qui j'ai adoré enseigner.

Merci à Cubi, Nikko, Béber, les M&Ms : Marianne et Major, Khroute et Joloss, Roumy. Un gros bisou à Sabrina et David. Je salue mes 'anciens' professeurs, Mme Samain, M. Voedts et M. Juhel, d'un constant soutien.

Un bisou énorme à Monsieur Marc Develter, qui malgré la distance est ma meilleure béquille depuis si longtemps déjà. Promis maintenant que la thèse est pliée, je remplis une valise de bière belges et je débarque en Italie.

Un grand coucou aux carnavaleux, qui m'ont apporté ma bouffée d'air annuelle nécessaire, Tante Lyne, Papy le Jardinier, Rougita, Maud, Dame Nature (Mono Têt'ch), Angelina et Grand Daniel, Val et Mimile, tous les Reutelaeres... et bien sûr mon corsaire.

Enfin je remercie toute ma famille et en particulier Maman, Papa, Simon, Lucie et Camille.

Je suis très fière de vous tous. Bravo et merci.

Lille, le 19 juillet 2010.

« Depuis qu'on a découvert
que deux plus deux font quatre
il n'y a plus grand chose à inventer
en mathématiques. »
Papy

Table des matières

Introduction générale **13**

1 Modélisation d'une population de métastases 15
1.1 Le modèle 15
1.2 Analyse mathématique 17
1.3 Simulations numériques 18
2 Modélisation de la respiration 18
2.1 Simulation du flux d'air dans les poumons 19
2.2 Modélisation des échanges gazeux 20
3 Modélisation d'une cascade protéinique 22
4 Modèle de Patlak–Keller–Segel 1D 23
4.1 Etude mathématique du système après explosion 23
4.2 Étude numérique 24
5 Plan de la thèse 24

Bibliographie **25**

Partie I Modèle pour une population de métastases

Chapitre 1
Un modèle décrivant la croissance et la distribution en taille d'une population de métastases
In *DCDS–B* Volume 12 No 4 (2009)

1.1 Introduction 29
1.2 Analysis of the PDE 31
1.2.1 Existence and uniqueness of a solution 31
1.2.2 A generational point of view 34
1.3 Eigenproblem 35

1.4 General relative entropy and asymptotic behavior 37
1.4.1 Conservation law and general relative entropy 37
1.4.2 Boundedness properties 38
1.4.3 Asymptotic behavior 38
1.5 Numerics 45
1.5.1 Numerical problems linked to the equation and the boundary condition 46
1.5.2 Numerical scheme 46
1.5.3 Numerical results 50
1.6 Conclusion and perspectives 54
1.7 Appendix : Discrete model 56
1.7.1 The discrete model 56
1.7.2 Estimates 58
1.7.3 Proof of Theorem 1.7.1 60
1.8 Appendix : Explicit computation of the solution using Laplace's transform 61

Bibliography **69**

Partie II Modélisation de la respiration

Chapitre 2
Le poumon humain

2.1 Introduction 73
2.2 Physiologie de la respiration 74
2.2.1 Mécanique de la ventilation 74
2.2.2 Géométrie des voies respiratoires 74
2.2.3 Physiologie des voies respiratoires 75
2.2.4 Les échanges gazeux 78
2.3 Mécanique de la respiration 80
2.3.1 Un peu d'hydrodynamique 80
2.4 Modèles couplés 83
2.4.1 Le modèle mécanique 0D 83
2.4.2 Un modèle multi–échelle pour l'appareil respiratoire 85
2.4.3 Un modèle pour la diffusion de l'oxygène dans l'arbre bronchique 89

Chapitre 3
Une méthode numérique pour la simulation du flux d'air dans les poumons

3.1 Introduction . . . 91
3.2 Analyse mathématique . . . 93
3.2.1 Formulation variationnelle . . . 93
3.2.2 Bilan d'énergie . . . 93
3.2.3 Réécriture du problème . . . 94
3.2.4 Existence et unicité d'une solution . . . 95
3.3 Méthode numérique . . . 96
3.3.1 Le système de Stokes . . . 98
3.3.2 Extension au problème de Navier–Stokes . . . 101
3.3.3 Méthodes de projection . . . 102
3.3.4 Pertinence du choix d'une méthode de projection . . . 104
3.4 Simulations numériques . . . 104
3.4.1 Premier modèle . . . 105
3.4.2 Respiration forcée . . . 105
3.4.3 Un cas de référence : données non pathologiques . . . 107
3.4.4 Étude de la sensitivité . . . 107
3.4.5 Simulations 3D . . . 113
3.5 Conclusions et perspectives . . . 117

Chapitre 4
Un modèle pour l'absorption de l'oxygène dans les poumons

4.1 Introduction . . . 119
4.2 Un modèle mono–compartiment pour l'oxygène . . . 121
4.3 Prise en compte de l'hétérogénéité : état de l'art . . . 123
4.4 Un modèle 1D . . . 125
4.4.1 Dérivation du modèle . . . 125
4.5 Simulations numériques . . . 129
4.5.1 Résultats numériques . . . 130
4.5.2 Discussion . . . 131
4.6 Conclusions et perspectives . . . 132
4.7 Appendix : Détails pour le calcul . . . 133
4.7.1 Calcul de $\partial_t s_b$ et $\partial_t s_a$ . . . 133
4.7.2 Calcul de ℓ_a . . . 133
4.7.3 Calcul de $W(\ell)$ . . . 134

Chapitre 5
Annexe : Données numériques

5.1 Données morphométriques de WEIBEL pour le poumon . . . 135
5.2 Paramètres pour le fluide . . . 136

5.2.1 Valeurs 3D . . . 136
5.2.2 Valeurs 2D . . . 136
5.3 Paramètres pour le système masse-ressort . . . 136
5.4 Modèle non–linéaire : constante de raideur . . . 136
5.5 Volumes bronchiques . . . 137
5.6 Trois scénarii pour la force extérieure . . . 137
5.6.1 Au repos . . . 137
5.6.2 A l'exercice . . . 137
5.6.3 Manoeuvre forcée . . . 137
5.7 Volumes d'oxygène absobé . . . 138
5.8 Ratio ventilation/perfusion . . . 138

Bibliographie **139**

Partie III Modélisation d'une cascade protéinique

Chapitre 6
Propagation d'un signal dans la cascade MAPK : rôle de la bistabilité et de l'ultrasensitivité pour un problème mixte

6.1 Introduction and biological context . . . 145
6.1.1 The cell cycle . . . 146
6.1.2 *Xenopus* oocyte . . . 146
6.2 Formulation of the problem . . . 147
6.3 Dimensionless equations . . . 151
6.4 Stability and wave propagation . . . 152
6.4.1 Preliminary remarks . . . 152
6.4.2 Study of the pure initial value problem on the half–plane $\mathbb{R} \times \mathbb{R}^+$ . . . 153
6.4.3 Study of the initial-boundary value problem (6.14)–(6.16) in the quarter–space $\mathbb{R}^+ \times \mathbb{R}^+$ 162
6.5 Numerical Study of the Wave Propagation . . . 170
6.5.1 Approximation of the solution . . . 170
6.5.2 Asymptotic velocity . . . 177
6.5.3 Influence of the Hill exponent . . . 178
6.6 Conclusion and perspectives . . . 180
6.7 Appendix A : Derivation of the expression of the reaction term $\Sigma[m(x,t)]$. . . . 181
6.8 Appendix B : Properties of q_β defined at (6.29) . . . 183

Bibliography **187**

Partie IV Modèle de Patlak–Keller–Segel 1D

Chapitre 7
Description et simulation du système de Patlak–Keller–Segel modifié 1D après « blow–up »

7.1 Introduction . . . 191
7.1.1 Le modèle parabolique de Patlak– Keller–Segel . . . 192
7.1.2 Le modèle de Patlak–Keller–Segel modifié . . . 193
7.2 Description du système après « blow–up » . . . 194
7.2.1 Mesures de défaut . . . 194
7.2.2 Convergence de mesures : rappels et notations . . . 195
7.2.3 Estimations *a priori* et mesure de défaut diagonale . . . 196
7.2.4 Solutions mesures généralisées . . . 200
7.2.5 Système après blow–up . . . 202
7.3 Le schéma de Jordan–Kinderlehrer–Otto . . . 206
7.3.1 Philosophie . . . 207
7.3.2 Quelques rappels : distance de Wasserstein et transport optimal . . . 212
7.3.3 Écriture du schéma pour le système de Patlak–Keller–Segel (7.3) . . . 213
7.3.4 Cas particulier de la dimension 1 . . . 214
7.4 Adaptation du schéma après blow–up . . . 216
7.4.1 Heuristique . . . 217
7.4.2 Algorithme « cut & paste » . . . 217
7.4.3 Illustrations numériques . . . 222
7.5 Conclusion et perspectives . . . 224

Bibliographie **225**

Introduction générale

Les systèmes biologiques offrent une grande variété de phénomènes dont la plupart sont encore de nos jours méconnus et mal compris. Il existe de multiples raisons à cela.

La première est sans doute la multitude ; chaque pathologie, chaque phénomène biologique constitue un sujet d'étude à lui tout seul, ce qui préserve encore pour longtemps les scientifiques du désoeuvrement.

La deuxième est sans doute la grande complexité de la plupart de ces phénomènes. Citons par exemple le cas des cascades protéiniques, où une même protéine, telle MAPK qui sera l'objet d'une partie des travaux de cette thèse, peut avoir de multiples rôles et où un même effet peut avoir des causes multiples : par exemple le franchissement de la transition G2/M et l'entrée en phase M, lors du cycle cellulaire des ovocytes de xénopes[1] peuvent être obtenus par la sur-expression de MEK (MAPK Kinase) ou bien encore par stimulation distincte de deux voies de signalisations différentes(voir [Rus09]).

La dernière que nous citerons est la raison essentielle qui motive la modélisation mathématique des phénomènes biologiques. Il s'agit des difficultés techniques et matérielles ; les limites de l'imagerie médicale : actuellement on ne peut déceler des tumeurs cancéreuses trop petites (de l'ordre de quelques mm^3) car elles sont invisibles à l'imagerie. Nous ne pouvons pas non plus visualiser ni séquencer les plus petites bronches d'un poumon (en pratique, seules les 10 premières générations sont visualisables). D'autre part, l'expérimentation est parfois impossible : il est impossible par exemple d'introduire un instrument de mesure dans le poumon sans perturber le bon fonctionnement de celui–ci et donc fausser la mesure. Certaines expérimentations ne sont jamais reproductibles dans les mêmes conditions, pensons par exemple aux tests cliniques pour des patients atteints de cancers ou tout simplement interdites pour des raisons éthiques. Certaines expérimentations envisageables, sont parfois très difficiles à exécuter même *in vitro*, par exemple les ovocytes de xénopes sont opaques, ce qui rend très difficile l'expérimentation et nécessite l'utilisation de marqueurs particuliers. Leurs coûts et les temps d'expérimentation sont aussi des contraintes non négligeables. La modélisation mathématique peut devenir alors un outil puissant pour le biologiste : l'expérimentation *in silico*, c'est–à–dire dans l'ordinateur, reproductible à souhait et peu coûteuse multiplie les possibilités et permet de mieux préparer et envisager l'expérimentation *in vitro*, voire *in vivo*.

Cette thèse illustre la diversité des problèmes biologiques pour lesquels l'outil mathématique et la modélisation peuvent aider à mieux comprendre certains mécanismes ou pathologies. En effet, les travaux rassemblés ici touchent à quatre domaines tout à fait indépendants.

Une première partie est dédiée à l'étude d'une population de métastases. Une métastase est la croissance d'un organisme pathogène ou, dans notre cas ici, d'une cellule tumorale à distance du site initialement atteint. Les métastases font toute la gravité du cancer. En effet, l'extirpation chirurgicale complète d'un cancer ne permet pas d'avoir la certitude qu'une métastase ne se découvrira pas des mois ou des années plus tard. Il y a des métastases qu'on ne peut pas voir à l'imagerie médicale, et qui donneront lieu à de nouvelles tumeurs : c'est la rechute. De plus, de tels modèles sont un outil intéressant en pharmacologie, afin de tester de nombreuses stratégies thérapeutiques sans aucune incidence pour le patient. Cela est d'autant plus intéressant étant donné la toxicité des drogues employées dans ce genre de traitement.

La deuxième partie de ce travail est consacrée à l'étude et à la modélisation de la respiration. Comme nous l'avons mentionné, l'étude de la respiration et du poumon est rendue difficile par la structure particulière de cet organe ; les voies respiratoires et leur structure en arbre dichotomique confère au poumon une géométrie fractale : les petites bronchioles aux extrémités de l'arbre ont des tailles de l'ordre de 0.29 mm de diamètre [WSF05] rendant impossible leur visualisation directe par imagerie. De plus, dû à la géométrie en chou–fleur, deux points très

[1] Grenouille.

proches l'un de l'autre du point de vue de la distance euclidienne peuvent être très éloignés d'un point de vue distance parcourue dans l'arbre. Dans ce travail, nous nous sommes intéressés en particulier à deux aspects : la description et la simulation des flux d'air dans le poumon et la modélisation des échanges gazeux.

La troisième partie concerne la modélisation d'une cascade protéinique, c'est–à–dire une suite de réactions qui entraîne l'activation de certaines protéines ou la transmission d'un signal. Ces systèmes sont très complexes car l'expression d'une même protéine peut avoir de multiples effets. La cascade MAPK chez les ovocytes de xénopes, que nous étudierons en particulier dans cette thèse, peut être isolée par les biologistes et donc constitue un point de départ intéressant pour l'étude de ce genre de cascade. Ce travail peut être considéré comme une brique de base pour la modélisation de systèmes plus grands. À plus long terme, les MAPK étant impliquées dans les mécanismes de multiplication cellulaire, une meilleure compréhension de la cascade pourrait déboucher sur des thérapies pour les pathologies affectant le cycle cellulaire, typiquement les cancers.

Enfin, notre quatrième partie concerne encore un autre phénomène biologique : il s'agit des phénomènes de chimiotactisme. Le chimiotactisme est le phénomène par lequel des cellules ou de petits organismes dirigent leurs mouvements en fonction de la concentration de certaines espèces chimiques présentes dans le milieu. Le chimiotactisme est omni–présent en biologie : trouver sa nourriture ou bien pour les phénomènes angiogénique, notamment lors de la croissance tumorale. Le modèle de Patlak–Keller–Segel qui va nous intéresser ici, concerne l'auto–organisation, à savoir le signal chimique est émis par les organismes eux–mêmes. De tels systèmes font apparaître comme nous le verrons, des points d'agrégation.

Avant de rentrer plus dans les détails, notons qu'un point commun à tous les modèles présentés dans ce manuscrit sont leur simplicité que l'on mesure ici à l'aide du nombre de paramètres. En effet, nous sommes convaincus que pour comprendre l'influence de chacun des paramètres, ceux–ci doivent être en nombre réduit. Cette volonté de simplicité conduit à des choix de modélisation qui se doivent pertinents.

Apports du point de vue de la modélisation

Certains des thèmes abordés reprennent des modèles déjà existants, proposés par des biologistes ou physiciens, et le travail effectué consiste essentiellement en l'analyse mathématique et/ou numérique de ceux–ci. Pour d'autres, le travail a été essentiellement un travail de modélisation. Ceci est en particulier le cas pour la modélisation des échanges gazeux dans le poumon et la modélisation de la cascade protéinique MAPK.

L'hétérogénéité des échanges gazeux dans le poumon avait déjà été décrite par de nombreux auteurs [Pai73, SFW02, WSF05, HGC$^+$05]. Des modèles ont été proposés, notamment des modèles 'labyrinthes' [FFS$^+$05a, FFS05b, SFW02], mais ils ne concernent cependant que la diffusion de l'oxygène dans l'acinus[2]. L'originalité de notre modèle est de considèrer l'appareil respiratoire dans son ensemble, modélisant ainsi à la fois la convection et la diffusion de l'oxygène dans l'arbre bronchique et liant directement la force exercée par le diaphragme lors de la respiration au volume d'oxygène absorbé.

L'autre apport essentiel d'un point de vue de la modélisation de cette thèse concerne la modélisation de la cascade protéinique MAPK, plus particulièrement la propagation de front d'onde pour cette cascade. Encore une fois, de nombreux travaux ont déjà été effectués sur ce sujet [AFS04, Kho00, MnGNK09, RGBB09], mais une des originalités de ce travail est la prise en compte des effets spatiaux.

Apports mathématiques

Les applications à la biologie peuvent faire émerger de nouveaux problèmes et questions mathématiques et ce travail en est la confirmation.

L'équation obtenue pour la modélisation d'une population de métastases fait partie d'une famille plus large d'équations aux dérivées partielles décrivant une population structurée appelées équations de McKendrick–Von Foerster. L'étude de telles équations a été développée notamment par MICHEL [Mic05] qui utilise comme outil la notion d'entropie relative généralisée. Dans notre cas, l'annulation de la vitesse apporte des difficultés supplémentaires, notamment pour montrer la compacité de l'opérateur. L'apport de notre travail consiste essentiellement dans le contournement de cette difficulté.

D'autre part, le modèle multi–échelle utilisé pour la description des flux d'air dans le poumon, conduit à un système de Navier–Stokes muni de conditions aux bords dissipatives non classiques. L'inconvénient de telles

[2]Grappe d'alvéoles pulmonaires, lieu de l'échange gazeux dans le poumon

conditions est qu'elles ne sont pas facilement implémentables dans les logiciels éléments finis, tels `Freefem++` [HLHOP], à moins de rentrer en profondeur dans le code. Nous proposons une méthode qui a l'avantage de palier à cette difficulté. En effet, elle réduit le problème à la résolution de problèmes de Stokes avec conditions aux bords de Dirichlet–Neumann classiques, qui peuvent quant à eux être très facilement codés à l'aide de tout logiciel éléments finis classique.

Enfin, concernant le système de Patlak–Keller–Segel, il est connu qu'il existe un phénomène de masse critique au delà de laquelle il y a explosion en temps fini. Pour l'étude de ce système, l'intérêt de l'interprétation en termes de flot gradient est démontrée [BCC08]. Le schéma numérique proposé dans ce travail, fondé sur cette idée de flot gradient permet de simuler, pour le cas 1D du moins, le système avant explosion. Nous étendons ce travail aux temps après explosion d'une part en donnant un cadre mathématique pour la description du système après explosion, fondée sur les idées de mesure de défaut développée par POUPAUD [Pou02] et d'autre part, en étendant le schéma numérique aux temps après explosion à l'aide d'un algorithme mettant en évidence des phénomènes assez fins.

Dans ce qui suit, nous résumons, thème par thème, les résultats apportés par cette thèse. Nous renvoyons à l'introduction de chaque partie pour une description de chaque thème dans son contexte biologique et mathématique.

1 Modélisation d'une population de métastases

Le cancer est devenu ces dernières années la première cause de mortalité dans les pays occidentaux. Cette partie du travail, en collaboration avec PAULINE LAFITTE et THIERRY GOUDON, concerne la modélisation d'une population de métastases. Ce modèle avait déjà été proposé par IWATA *et al.* en 2000 [IKS00] mais n'avait pas été étudié mathématiquement dans les détails. Ainsi, notre travail concerne essentiellement l'étude mathématique et la simulation numérique du système.

L'équation fournie par le modèle est une équation aux dérivées partielles décrivant une population structurée, en taille ici. Elle est complétée par une condition au bord non locale. Ce système fait partie de la famille des équations de type McKendrick–Von Foerster, qui ont été étudiées notamment par MICHEL [Mic05], en introduisant la notion d'entropie relative généralisée. Nous adapterons cette méthode à notre modèle. Dans notre cas, une difficulté est apportée par l'annulation de la vitesse en un certain point.

D'un point de vue numérique, nous menons une étude soignée, confortant l'étude théorique et pouvant servir de marche–pied pour traiter des modèles plus complexes. Nous verrons que la solution étant très raide notamment proche du bord $x = 1$, une méthode d'ordre élevée est utile ; nous utiliserons alors le schéma WENO (Weighted Essentially Non–Oscillatory).

Soulignons que des études similaires ont été effectuées par l'équipe de BARBOLOSI, BENABDALLAH, HUBERT, VERGA et BENZEKRY, cependant les outils utilisés sont différents tant sur le plan de l'analyse mathématique que sur celui de la simulation numérique.

1.1 Le modèle

Nous considérons une population de métastases. Chacune des métastases est caractérisée par sa taille notée $x \geq 1$, qui est à comprendre comme le nombre de cellule qui la constitue. Nous nous intéressons au comportement de la distribution en taille de cette population $\rho(t, x)$, où

$$\int_u^v \rho(t, x)\, \mathrm{d}x$$

donne le nombre de tumeurs dont la taille est comprise entre u et v à l'instant t. La dynamique est dirigée par deux phénomènes :

- ⋆ une tumeur de taille x croît à la vitesse $w(x) \geq 0$;
- ⋆ une tumeur émet des cellules cancéreuse au taux $\beta(x) \geq 1$.

Le système obtenu est le suivant :

$$\partial_t \rho + \partial_x (w\rho) = 0, \tag{1a}$$

$$w(1)\rho(t,1) = \int_1^\infty \beta(y)\rho(t,y)\,\mathrm{d}y + \beta(x_p(t)), \tag{1b}$$

La deuxième équation (1b), correspond à l'ensemble des tumeurs de taille 1 (cellules) produites à chaque instant. On distingue celles émises par la tumeur primaire $\beta(x_p(t))$, où x_p est la taille de la tumeur primaire solution de l'EDO :

$$\frac{\mathrm{d}}{\mathrm{d}t} x_p(t) = w(x_p(t)), \quad x_p(0) = 1,$$

et celles émises par toutes les autres tumeurs, données par le terme intégral. Le problème est complété par une condition initiale ; nous supposons qu'initialement il n'y a pas de métastase, ainsi

$$\rho\big|_{t=0} = 0.$$

Les hypothèses que nous imposons sur les fonctions w et β sont les suivantes :

$$\begin{cases} w \in C^1(\mathbb{R}), \\ |w(x) - w(y)| \leq L|x-y|, \\ w(x) > 0 \text{ pour tout } x \in [1,b[,\ w(b) = 0, \end{cases} \tag{2}$$

et

$$\beta \in L^\infty([1,b]),\ \beta(x) > 0 \text{ sur } [1,b]. \tag{3}$$

Le paramètre b représente la taille maximale des tumeurs. La particularité de ce système est que la fonction w s'annule en b et en particulier

$$\int_1^b \frac{\mathrm{d}x}{w(x)} = \infty. \tag{4}$$

En comparaison aux équations de McKendrick–Von Foerster, où typiquement $w \equiv 1$, cette singularité induit des difficultés mathématiques.

En pratique, nous choisissons pour w le taux de Gompertz

$$w(x) = ax\ln(b/x), \tag{5}$$

et pour β

$$\beta(x) = mx^\alpha. \tag{6}$$

Nous verrons que l'essentiel de l'information sur la croissance de la population est concentrée dans la valeur d'un paramètre positif λ_0 caractérisant la croissance de la population et appelé dans le contexte de dynamique des population taux Malthusien. En effet, on montre que

$$\rho(t,x) \sim Ce^{\lambda_0 t} N(x),$$

lorsque t tend vers l'infini. Nous verrons que λ_0 et N sont les éléments propres du système. La résolution de ce problème aux valeurs propres se fait habituellement en utilisant le théorème de Krein–Rutman, équivalent du théorème de Perron–Frobenius en dimension infinie. Cependant, la singularité de la fonction w rend difficile la preuve directe de la compacité de l'opérateur nécessaire à l'application du théorème de Krein–Rutman. La difficulté est contournée par l'utilisation d'un changement de variable astucieux qui simplifie grandement la question. Enfin, le comportement asymptotique est déterminé par l'utilisation de la méthode d'entropie.

D'un point de vue numérique, nous avons dû tenir compte de deux contraintes :

- ⋆ nous voulons simuler le système en temps grand afin de capter l'asymptotique, ce qui réclame une attention particulière vis–à–vis du schéma numérique choisi ;
- ⋆ le domaine de calcul est très grand ($b = 10^{11}$), ce qui rend impossible l'utilisation de maillage très fins.

Nous allons à présent énoncer les principaux résultats obtenus pour ce modèle, tant sur le plan de l'analyse mathématique que sur celui du numérique.

1.2 Analyse mathématique

Nous montrons que, sous les hypothèses (2)–(3), il y a existence et unicité d'une solution mesure au problème (1) (Théorème 1.2.3). Le cadre mathématique étant précisé à la Définition 1.2.2.

L'étude mathématique effectuée ensuite se concentre sur le problème où nous enlevons le terme source en $\beta(x_p(t))$:

$$\begin{aligned} &\partial_t \rho + \partial_x(w\rho) = 0, \\ &w\rho(t,1) = \int_1^b \beta(y)\rho(t,\mathrm{d}y), \\ &\rho\big|_{t=0} = f_0, \end{aligned} \tag{7}$$

Ce problème est équivalent au précédent ; en effet, nous remarquons que si ρ est solution du problème original (1), alors $f = \rho + \delta(x = x_p(t))$ est solution de (7) avec $f_0 = \delta(x = 1)$. Ainsi les deux problèmes sont équivalents.

Nous étudions ensuite le comportement asymptotique de la solution. L'outil utilisé est celui développé par Michel [Mic05], à savoir la notion d'entropie relative généralisée. Une étape préliminaire à cette méthode est l'étude du problème aux vecteurs propres associé au système. Nous cherchons (λ_0, N, Φ) vérifiant

$$\frac{\partial}{\partial x}\left(N(x)w(x)\right) + \lambda_0 N(x) = 0, \tag{8a}$$

$$w(1)N(1) = \int_1^b \beta(y)N(y)\,\mathrm{d}y, \tag{8b}$$

$$-w(x)\frac{\partial}{\partial x}\Phi(x) + \lambda_0\Phi(x) = \Phi(1)\beta(x), \tag{8c}$$

$$\lambda_0 > 0, \quad N \geq 0, \quad \Phi \geq 0, \quad \int_1^b N\Phi\,\mathrm{d}x = 1, \quad \int_1^b N\,\mathrm{d}x = 1. \tag{8d}$$

Ce système admet une unique solution. En effet, nous montrons :

Théorème 0.1.1 *Il existe un unique triplet $\lambda_0 > 0$ et $(N, \Phi) \in L^1([1,b]) \times L^\infty([1,b])$ vérifiant* (8a)-(8d).

Ce genre de théorème est classiquement démontré en utilisant le théorème de Krein–Rutman, version de Perron–Frobenius pour la dimension infinie. Cependant, ici, dû à la singularité de la fonction w (voir(4)), la compacité de l'opérateur ne peut être obtenue facilement. La méthode que nous employons est fondée sur un changement de variable, permettant de montrer directement l'existence et l'unicité du triplet.

Une fois l'existence d'un tel triplet d'éléments propres obtenue, le théorème suivant montrant la décroissance de l'entropie, permet de donner des bornes pour la solution dans diverses normes et enfin de donner le comportement asymptotique de la solution.

Théorème 0.1.2 *Soit (N, λ_0, Φ) le triplet solution du problème* (8a)-(8d)*, alors pour tout $\rho(0,\cdot) \geq 0$ et tout $\rho(t,\cdot)$ solution de* (7) *avec condition initiale $\rho(0,\cdot) \in L^\infty\left([1,b], \frac{\mathrm{d}x}{N}\right)$,*

1. *nous avons la loi de conservation*

$$\int_1^b \Phi(x)\rho(t,x)e^{-\lambda_0 t}\,\mathrm{d}x = \int_1^b \Phi(x)\rho(0,x)\,\mathrm{d}x,$$

2. *et pour toute fonction convexe H, l'entropie relative généralisée définie par*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_1^b \Phi(x)N(x)H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}x := -D_H,$$

avec

$$\begin{aligned} -D_H \;=\; & w(1)N(1)\Phi(1) \\ & \times\left[H\left(\int_1^b \frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\,\mathrm{d}\mu(x)\right) - \int_1^b H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}\mu(x)\right], \end{aligned}$$

où

$$\mathrm{d}\mu(x) = \frac{\beta(x)N(x)}{w(1)N(1)}\,\mathrm{d}x,$$

négatif.

Nous obtenons alors le théorème suivant décrivant le comportement en temps long de la solution :

Théorème 0.1.3 *Pour tout* $\rho(0,\cdot) \geq 0$ *tel que* $\frac{\rho(0,\cdot)}{N(\cdot)} \in L^\infty \cap L^2([0,b], N\Phi\,\mathrm{d}x)$ *et pour tout* $\rho(t,\cdot)$ *solution de* (7) *avec condition initiale* $\rho(0,\cdot)$ *nous avons la convergence suivante*

$$\frac{\rho(\cdot,t)e^{-\lambda_0 t}}{N(\cdot)} \xrightarrow[t\to\infty]{} \int_1^b \Phi(x)\rho(0,x)\,\mathrm{d}x,$$

dans $L^\infty \cap L^2([0,b], N\Phi\,\mathrm{d}x)$ *lorsque* t *tend vers l'infini.*

La preuve de ce théorème s'appuie sur les propriétés de décroissance de l'entropie données par le théorème précédent. Cependant, une attention particulière doit être donnée aux passages à la limite. Nous complétons ce résultat en précisant la vitesse de convergence (Théorème 1.41) sous quelques hypothèses supplémentaires (voir page 43 pour plus de détails), vérifiées dans le cas particulier des choix effectués pour les fonctions w (5) et β (6).

1.3 Simulations numériques

Comme mentionné précédemment, plusieurs contraintes sont à prendre en compte dans l'étude numérique du système (1) :

- ⋆ sans une discrétisation fine du domaine, ce qui n'est pas envisageable ici ($b = 7.3 \cdot 10^{10}$), la solution calculée de façon naïve explose rapidement ;
- ⋆ dans le même ordre d'idée, la condition au bord peut également devenir très grande, comparée à la valeur moyenne de la solution sur le reste du domaine ;
- ⋆ enfin, nous voulons observer l'asymptotique de la solution, pour cela nous avons besoin d'effectuer des calculs en temps long.

Notre stratégie est la suivante,

i) nous posons le changement de variable

$$y = \ln \frac{b}{x},$$

et de fonction

$$v(y,t) = be^{-y}u(y,t).$$

Ces changements permettent de réduire le domaine de calcul, tout en gardant la structure de l'équation (loi de conservation). De plus, la grille obtenue est raffinée proche du bord $x = 1$ (voir Figure 1.1), là où la solution est la plus raide, permettant ainsi de capturer au mieux l'influence de la condition au bord ;

ii) nous utilisons un schéma d'ordre élevé, WENO–5. Ce choix est essentiellement motivé par la différence d'échelle entre la condition au bord et la solution aux temps petits.

La résolution numérique permet de retrouver le comportement asymptotique de la solution, notamment le taux Malthusien (voir Figure 1.3).

Cette étude est enfin complétée par l'étude d'un modèle discret associé au modèle continu présenté et par le calcul explicite d'une solution du problème (1) par transformée de Laplace.

Ce modèle ainsi posé est très simple et décrit une population de métastases non–traitée. Le choix du taux de Gompertz pour la fonction w est sujet à controverses, et il y a sans doute encore des travaux de modélisation à réaliser afin d'améliorer ce choix. Ensuite, l'intérêt d'un tel modèle est de pouvoir intégrer des stratégies thérapeutiques telles que par exemple des traitements anti–angiogénique qui modifieraient les fonctions de croissance w et d'émission β. Des travaux dans ce sens sont en cours au sein de l'équipe Marseillaise de BARBOLOSI *et al.* [Ben].

2 Modélisation de la respiration

Notre travail concernant la modélisation de la respiration à fait l'objet d'une étroite collaboration avec l'EPI REO de l'INRIA Rocquencourt. Ce travail comporte deux parties différentes : la première concerne la simulation du flux d'air dans l'arbre bronchique et la seconde la modélisation des échanges gazeux.

2.1 Simulation du flux d'air dans les poumons

Ce travail est le fruit d'une collaboration avec CÉLINE GRANDMONT, BÉRÉNICE GREC, BERTRAND MAURY et DRISS YAKOUBI. Le modèle utilisé dans cette partie a déjà été décrit dans [BGM09], et le travail effectué consiste essentiellement en la mise en place d'une méthode numérique.

Le modèle

L'appareil respiratoire a une géométrie complexe : pour un calcul direct, c'est un arbre dichotomique composé d'en moyenne 23 générations qui devrait être maillé. Cependant ceci est infaisable numériquement. De plus, jusqu'à présent, les bronches distales situées à partir de la neuvième génération ne sont pas visualisables, ni segmentables par le matériel médical d'imagerie usuel. Afin de simuler les flux d'air dans ces parties de poumon, nous avons donc besoin de modèles mathématiques simples, mais réalistes.

Le modèle que nous utilisons répond à ces critères. Il distingue trois parties dans le poumon :

- la partie proximale (générations 0 à 7–9), où nous supposons que les équations de Navier–Stokes sont vérifiées ;
- la partie distale (générations 8–10 à 17), où nous supposons que la loi de Poiseuille est vérifiée, et donc où chaque sous–arbre peut être représenté par un tube dont la résistance R_i est équivalente à celle du sous–arbre correspondant.
- les acini nichés dans le parenchyme pulmonaire. Nous décrirons le mouvement de celui–ci par un simple modèle mécanique 0D.

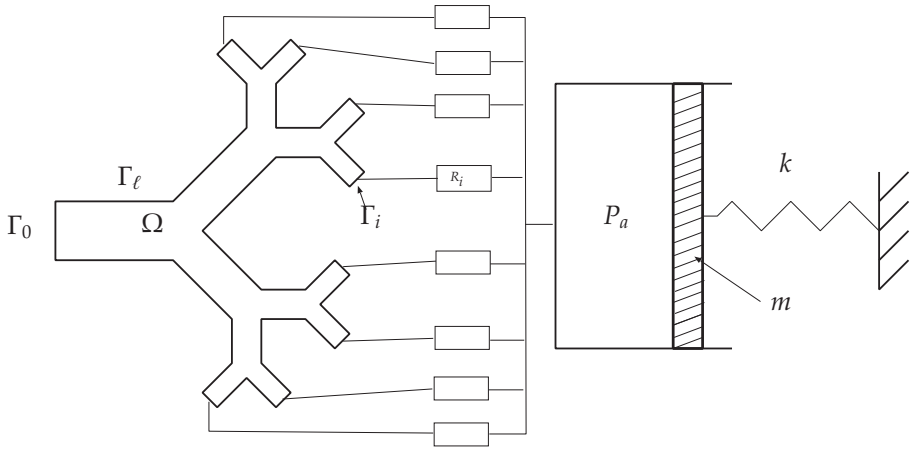


FIG. 1: Modèle multi–échelle

La Figure 1 illustre de manière synthétique la décomposition multi–échelle considérée. Ce modèle aboutit au système couplé suivant :

$$
\begin{cases}
\rho\dfrac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \eta\triangle\mathbf{u} + \nabla p &= 0\,, & \text{sur } (0,T)\times\Omega\,,\\
\nabla\cdot\mathbf{u} &= 0\,, & \text{sur } (0,T)\times\Omega\,,\\
\mathbf{u} &= 0\,, & \text{sur } (0,T)\times\Gamma_\ell\,,\\
\eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &= -P_0\mathbf{n}\,, & \text{sur } (0,T)\times\Gamma_0\,,\\
\eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &= -P_a\mathbf{n} - R_i\left(\displaystyle\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n}\,, & \text{sur } (0,T)\times\Gamma_i\,,\ i=1\,,\ \ldots\,,\ N\,,\\
\\
m\ddot{x} + (RS^2+\mu)\dot{x} + kx &= f_{ext} + SP_a\,,\\
S\dot{x} &= \displaystyle\sum_{i=1}^{N}\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n} = -\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}.
\end{cases}
$$

Les inconnues sont la vitesse et la pression du fluide, $\mathbf{u}$ et p ainsi que la pression alvéolaire P_a et le déplacement du diaphragme x. La force f_{ext} désigne les efforts exercés pour déplacer le diaphragme et représente le "moteur" de la respiration. Ce système doit être complété par des conditions initiales adaptées

$$\begin{cases} (\mathbf{u}, x, \dot{x})|_{t=0} = (\mathbf{u}_0, x_0, x_1), \text{ avec} \\ \nabla \cdot \mathbf{u}_0 = 0\,, \ \mathbf{u}_0 = 0 \text{ sur } \Gamma_\ell\,, \ \text{ et } Sx_1 = -\displaystyle\int_{\Gamma_0} \mathbf{u}_0 \cdot \mathbf{n}. \end{cases}$$

Une particularité de ce système est que toutes les sorties $\Gamma_i, 1 \leq i \leq N$ sont couplées (à travers le modèle 0D). De plus, les conditions aux limites sur les frontières artificielles Γ_i sont non standards puisqu'elles font intervenir les flux des vitesses fluides.

L'étude d'existence de solutions de ce problème a déjà été réalisée notamment dans [BGM09]. Dans ce travail, les auteurs proposent également une méthode numérique basée sur une discrétisation éléments finis du système qui implique, en particulier, la modification du produit matrice–vecteur associé au problème discrétisé, et ce afin de prendre en compte les conditions aux limites non standards. Cela nécessite d'intervenir en profondeur dans le code éléments finis utilisé. Nous avons donc voulu nous affranchir de cette contrainte. L'originalité de notre travail consiste en la décomposition du problème en plusieurs problèmes de Stokes avec conditions aux bords mixtes standards Dirichlet–Neumann ne nécessitant ainsi pas de modification du code pré–existant.

Méthode numérique

Après une discrétisation en temps du problème couplé et discrétisant la dérivée totale en temps de la vitesse du fluide par la méthode des caractéristiques, nous obtenons un système de Stokes muni de conditions aux bords non–standards mettant en jeu les flux d'air aux interfaces. Comme nous l'avons mentionné la difficulté consiste alors à traiter ces conditions dissipatives à l'aide d'un logiciel éléments finis par exemple `Freefem++` [HLHOP]. La stratégie est d'écrire la solution comme combinaison linéaire des solutions élémentaires pré–calculées et d'un terme correctif correspondant à la partie non-stationnaire du problème. Chacune de ces quantités est solution d'un problème de Stokes avec des conditions aux bords mixtes standards Dirichlet–Neumann, que l'on peut résoudre facilement à l'aide de tout logiciel éléments finis. Les coefficients de la combinaison linéaire sont ensuite calculés de façon à ce que les conditions aux bords dissipatives soient vérifiées. Nous renvoyons aux pages 99 et suivantes, pour une description plus détaillée de la méthode.

Cette méthode est ensuite complétée par une méthode de projection [Cho68, Cho69] pour la discrétisation en temps des équations de Stokes, permettant d'accélérer les calculs.

Les simulations numériques 2D réalisées permettent de retrouver les courbes débit–volume que les médecins sont habitués à lire (voir Figure 3.11) ; nous retrouvons à la fois le comportement qualitatif, les courbes ont la même allure, et les bons ordres de grandeur. Enfin, nous réalisons ensuite une étude numérique approfondie du modèle, afin de déterminer l'influence de chacun des paramètres, essayant à chaque fois de retrouver les caractéristiques de certaines pathologies pulmonaires. Nous terminons par donner quelques résultats 3D.

2.2 Modélisation des échanges gazeux

Ce travail est entièrement un travail de modélisation. Il est le fruit d'une collaboration avec SÉBASTIEN MARTIN et BERTRAND MAURY.

La géométrie particulière du poumon a pour conséquence l'hétérogénéité des échanges gazeux. En particulier, on observe l'importance de la géométrie lors du passage du régime de respiration au repos à celui de l'effort. Grossièrement, si on regarde les volumes d'oxygène absorbés au repos et à l'exercice, nous gagnons un facteur 9 (voit Table 5.4). Un facteur 3 est gagné par l'augmentation du rythme cardiaque (plus précisément, c'est ce que nous appellerons le ratio ventilation/perfusion qui augmente, voir page 138). Il nous manque donc un facteur 3. L'objectif de ce travail est de justifier que ce facteur 3 peut être obtenu grâce à la mécanique et à la géométrie.

Deux phénomènes principaux sont à prendre en compte :

- du fait de la géométrie, la vitesse du fluide décroît le long de l'arbre bronchique et passe au bout d'un certain temps en–dessous de la vitesse de diffusion de l'oxygène dans l'air. Cette transition a lieu approximativement à l'entrée des acinis, mais cela dépend fortement du régime auquel on se place ;

- tandis que l'oxygène diffuse le long de l'arbre, il est peu à peu absorbé par les alvéoles disposés le long de son parcours dans les voies aériennes, ce qui a pour conséquence l'appauvrissement de l'air en oxygène. C'est ce que nous appelons les effets d'écrantage.

La validation de notre modèle repose sur deux critères :

- quantitativement, nous voulons retrouver les volumes d'oxygène absorbé moyens par unité de temps à l'effort et au repos donnés dans la littérature de physiologie (voir Table 5.4) ;
- nous voulons observer les effets d'écrantage au repos : à savoir à l'exercice l'ensemble de la surface est disponible et donc l'oxygène est principalement absorbé par les dernières générations, tandis qu'au repos, la surface est minorée et l'oxygène est absorbé plutot par les générations 15 à 18 [WSF05].

Essentiellement, quatre paramètres entrent en ligne de compte pour le calcul des flux d'oxygène à travers la membrane pulmonaire :

1. la différence des pressions partielles en oxygène de chaque coté de la membrane, à savoir entre l'alvéole et le sang pulmonaire ;
2. la surface alvéolaire disponible ;
3. la perméabilité de la membrane, qui est une constante qui caractérise la facilité de l'oxygène à traverser la membrane ;
4. le ratio ventilation/perfusion.

Pour construire notre modèle, nous reprenons le modèle mécanique déjà utilisé pour la modélisation du mouvement du parenchyme et nous le couplons à un modèle 1D pour l'oxygène. Le poumon est alors représenté par un cylindre de rayon variable représentant les bronches (voir Figure 4.1) connecté à un réservoir représentant les alvéoles. Ce réservoir est lui même séparé du sang par une membrane poreuse (voir Figure 4.2). Nous obtenons le système d'équations suivant :

$$\dot{x} = u, \tag{9a}$$

$$\dot{u} = \frac{1}{m}\left(f_{ext} - \left(\frac{R_0 S^2}{1 + \theta S x / V_B^0} + \mu\right) u - kx\right), \tag{9b}$$

$$\partial_t s_b + \partial_\ell (s_b u^{\mathrm{f}}) = -\partial_t s_a, \tag{9c}$$

$$\partial_t (s_b c_b) + \partial_\ell (s_b c_b u^{\mathrm{f}}) + \partial_\ell (-D_{O_2} s_b \partial_\ell c_b) = -\partial_t s_a [\dot{x} \geq 0 \,?\, c_b, c_a] - W(\ell)\,(c_b - c_a), \tag{9d}$$

$$\partial_t (s_a c_a) = -b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}) + \partial_t s_a [\dot{x} \geq 0 \,?\, c_b, c_a] + W(\ell)\,(c_b - c_a). \tag{9e}$$

Les deux premières équations correspondent à la partie mécanique : x désigne le déplacement du parenchyme, u sa vitesse. f_{ext} est la force musculaire qui dirige la respiration. m est la masse déplacée, R_0 la résistance des bronches, μ la résistance des tissus et k la constante de raideur du ressort. S correspond à la surface de la partie mobile. Le rapport

$$\frac{1}{1 + \theta S x / V_B^0},$$

modélise l'activité du muscle lisse (voir Section 2.4.1, page 85, pour plus de détails). L'équation (9c), traduit la conservation du volume d'air ; s_b désigne la section bronchique et s_a la section équivalente alvéolaire, de sorte que $s_a(\ell)\,\mathrm{d}\ell$ donne le volume alvéolaire compris entre ℓ et $\ell + \mathrm{d}\ell$, la vitesse du fluide est notée u^{f}. L'équation (9d) correspond au bilan de masse pour l'oxygène écrit pour les bronches, tandis que la dernière équation (9e) correspond au bilan de masse écrit pour les alvéoles. c_a (resp. c_b) désigne la concentration en oxygène dans les alvéoles (resp. dans les bronches). Dans le membre de gauche de (9d), nous reconnaissons les termes de transport à la vitesse du fluide et de diffusion de l'oxygène.

Le terme commun aux équations (9d) et (9e) ;

$$\partial_t s_a [\dot{x} \geq 0 \,?\, c_b, c_a] + W(\ell)\,(c_b - c_a),$$

correspond aux termes d'échange entre bronche et alvéole ; le terme $\partial_t s_a[\dot{x} \geq 0 \,?\, c_b, c_a]$ est un terme convectif correspondant au volume d'oxygène entrant dans l'alvéole dû à la dépression causée par la déformation mécanique de l'alvéole (à l'inspiration) ou sortant de l'alvéole dû à la surpression causée cette fois par la compression de l'alvéole (à l'expiration). Le terme $W(\ell)\,(c_b - c_a)$ est un terme de diffusion dans la direction 'orthogonale' entre l'alvéole et la bronche.

Le terme de perte dans la troisième équation, Équation (9e),

$$-b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}),$$

correspond à la quantité d'oxygène absorbée qui passe dans le sang avec b la perméabilité, $\rho_{V/P}$ le ratio ventilation/perfusion, ℓ_a la densité de surface alvéolaire et $(P_0 c_a - P_{sang})$ la différence de pressions partielles en oxygène entre l'alvéole et le sang.

Nous réalisons ensuite des simulations numériques de ce système en prenant garde à la diffusion numérique. En effet, la valeur du coefficient de diffusion de l'oxygène dans l'air est très petit, $D_{O_2} = 0.2 \cdot 10^{-4}\mathrm{m}^2{\cdot}\mathrm{s}^{-1}$, et il faut prendre garde à ce que le coefficient de diffusion numérique ne soit pas supérieur à cette valeur. Ceci est d'autant plus important que nous cherchons à observer des phénomènes pour lesquels la diffusion de l'oxygène a un rôle primordial.

3 Modélisation d'une cascade protéinique

Ce travail, fruit d'une collaboration avec RALF BLOSSEY de l'IRI d'une part et PAULINE LAFITTE et THIERRY GOUDON d'autre part, propose un modèle pour décrire la voie de signalisation MAPK pour les ovocytes de xénopes. L'originalité de ce travail réside dans le modèle proposé pour lequel l'analyse mathématique, bien qu'elle repose sur des outils bien connus, permet d'exhiber des comportements remarquables et parfois inattendus.

La cascade MAPK est constituée de multiples réactions de phosphorylation (voir Figure 6.4 pour une représentation schématique). La cinétique de toutes ces équations n'étant pas la même, nous nous plaçons à l'équilibre pour les réactions rapides. Cette hypothèse, classique pour les chimistes, permet de simplifier grandement le problème. Toute la cinétique de la réaction est alors dirigée par la concentration en Mos, qui est la protéine à l'entrée de la cascade. De plus, nous supposons que la diffusion des protéines impliquées ne fait pas intervenir de structure particulière telle le squelette cytoplasmique.

Après adimensionnement, le système obtenu est le suivant :

$$\begin{array}{lll} \partial_t m - \partial_{xx} m = z_\alpha(m) - m & \text{sur} & \mathbb{R}^+ \times [0, L], \\ m(0, x) = m_0(x) & \text{sur} & [0, L], \\ m(t, 0) = \psi(t) & \text{sur} & \mathbb{R}^+, \end{array}$$

où

$$z_\alpha(m) = \frac{\alpha m^4}{m^4 + m^2 + 1} \tag{10}$$

correspond au terme de réaction. Remarquons, que sous cette forme, toute l'information est contenue dans le paramètre α. Un autre paramètre important d'un point de vue biologique est l'exposant de Hill. Biologiquement, il est lié à la notion d'ultrasensitivité de la cascade ; d'un point de vue mathématique, il correspond au degré des polynômes au numérateur et au dénominateur du terme non–linéaire de réaction (10).

L'étude mathématique de ces systèmes est bien connue et ce depuis longtemps [AW75]. Elle montre l'existence d'un front d'onde et d'une vitesse asymptotique associée. Le front d'onde n'apparaît que si la condition au bord et/ou la condition initiale apportent assez de masse au système. On observe un phénomène de seuil. Cependant, l'analyse mathématique ne permet pas d'expliciter ce seuil par une condition nécessaire et suffisante ; nous obtenons uniquement des conditions suffisantes (Théorèmes 6.4.10, 6.4.12,6.4.13, 6.4.14). De plus, dans notre cas, il n'existe pas de formule explicite pour la vitesse asymptotique.

Nous complétons donc l'étude mathématique par une étude numérique. Nous illustrons ce phénomène de seuil et nous calculons numériquement la vitesse asymptotique. Enfin nous étudions numériquement l'influence des paramètres sur le système à savoir α et l'exposant de Hill. Il s'avère que les résultats obtenus sont en contradiction avec l'intuition du biologiste.

4 Modèle de Patlak–Keller–Segel 1D

Cette dernière partie concerne la description et la simulation du système de Patlak–Keller–Segel 1D aux temps après explosion. Elle a fait l'objet d'une collaboration avec VINCENT CALVEZ et JOSÉ–ANTONIO CARRILLO. Le phénomène biologique dont il est question ici est le chimiotactisme. Plus particulièrement le système de Patlak–Keller–Segel concerne l'auto–organisation : nous considérons une population de cellules qui produit elle–même le signal chimique qui les attire.

Le système étudié est le suivant :

$$\begin{cases} \dfrac{\partial n}{\partial t} &= \Delta n - \nabla \cdot (\chi n \nabla c), \quad t > 0, x \in \Omega \subset \mathbb{R}^d, \\ c &= -\dfrac{1}{d\pi} \log |z| * n, \end{cases} \tag{11}$$

pour $d = 1$. Cette formulation, due à CALVEZ, PERTHAME et SHARIFI TABAR [CPSt], n'a aucune signification physique en dehors de la dimension $d = 2$, où le noyau logarithmique correspond à la formulation à l'aide du noyau de Poisson de la solution du Laplacien. Cependant, ces systèmes posés dans les différentes dimensions, ont l'avantage de partager de nombreuses propriétés (énergie libre, masse critique...). Numériquement, il est intéressant d'avoir une version unidimensionnelle du système de Patlak–Keller–Segel : les calculs 1D étant beaucoup moins coûteux, il est vraiment possible de réaliser des études numériques fines du modèle.

Comme mentionné précédemment, pour ce système posé en dimension $d = 1$ nous avons un phénomène de masse critique : si M la masse totale du système à l'instant initial vérifie $\chi M > 2\pi$ il y a explosion en temps fini, et des masses de Dirac apparaissent. L'objectif de ce travail est tout d'abord de décrire mathématiquement le système après explosion puis de le simuler numériquement.

4.1 Etude mathématique du système après explosion

La description du système après explosion que nous proposons fait appel à la notion de mesure de défaut introduite par POUPAUD [Pou02]. Une étude similaire a été réalisée par DOLBEAULT et SCHMEISER [DS09] pour le cas de la dimension $d = 2$. Dans notre cas, la convolution avec le noyau logarithmique ne correspond plus à la solution fondamentale du Laplacien et le gradient de la concentration est donné par l'opposé de la transformée de Hilbert. Cette perte de régularité apporte des difficultés supplémentaires, cependant nous obtenons que sous certaines hypothèses le système peut être décrit ainsi ; nous supposons

$$\rho = \bar{\rho} + \hat{\rho}$$

où $\bar{\rho}$ correspond à la partie régulière et $\hat{\rho}$ est une somme de masses de Dirac,

$$\hat{\rho}(t,x) = \sum_{\ell \in L} M_\ell(t) \delta_\ell(t,x)$$

avec $\delta_\ell(t,x) = \delta(x - x_\ell(t))$. x_ℓ correspond donc à la trajectoire de la ℓ^e masse de Dirac dont la masse est M_ℓ. Nous obtenons alors :

$$\begin{aligned} &\partial_t \overline{\rho} + \chi\, \partial_x(\overline{\rho}\ \partial_x S_0[\,\overline{\rho}\,]) - \sum_{\ell \in L} \frac{\chi}{\pi}\, \partial_x \left(\overline{\rho} \frac{M_\ell}{x - x_\ell} \right) - \Delta \overline{\rho} = 0, \\ &\dot{M}_\ell = \left(\frac{\chi}{\pi} M_\ell + 1 \right) \Big(\ \partial_x\, \overline{\rho}\,(t, x_\ell^+) - \ \partial_x\, \overline{\rho}\,(t, x_\ell^-) \Big), \\ &\dot{x}_\ell = \chi\, \partial_x S_0[\,\overline{\rho}\,](x_\ell) - \frac{\chi}{\pi} \sum_{m \neq \ell} \frac{M_m}{x_\ell - x_m}, \end{aligned}$$

où $S_0[\rho]$ correspond à la convolée de ρ avec le noyau logarithmique.

4.2 Étude numérique

Nous simulons ensuite ce système numériquement. Le schéma proposé prolonge celui de BLANCHET *et al.* [BCC08] aux temps après explosion. Ce schéma, fondé sur l'interprétation d'OTTO de l'équation en termes de flot gradient pour la métrique de Wasserstein, peut s'écrire facilement sur l'inverse de la fonction de répartition de notre densité de cellules. Pour cette fonction, un point d'accumulation de la densité correspond à un plateau. Pour étendre ce schéma aux temps après explosion, nous coupons les plateaux et les considérons comme masses de Dirac, nous recollons les deux branches restantes afin de reformer la partie continue, et nous gérons numériquement les interactions entre partie continue et masses de Dirac à l'aide de la description mathématique du système obtenue précédemment. La difficulté réside dans la méthode de découpage. Ce choix fait appel à une réflexion assez fine sur le système.

5 Plan de la thèse

Comme nous l'avons dit ce travail est composé de quatre parties bien distinctes et tout à fait indépendantes tant sur le plan des thématiques biologiques évoquées, que sur le plan mathématique. La lecture des différentes parties peut donc se faire totalement indépendemment les unes des autres. Le lecteur trouvera en début de chaque partie une introduction au contexte biologique et mathématique du travail correspondant, ainsi qu'en fin de partie, les perspectives et travaux futurs envisagés.

Bibliographie

[AFS04] David Angeli, James E. Ferrell, and Eduardo D. Sontag.
Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems.
Proceedings of the National Academy of Sciences of the United States of America, 101(7) :1822–1827, 2004.

[AW75] Donald G. Aronson and Hans F. Weinberger.
Nonlinear diffusion in population genetics, combustion, and nerve propagation.
Lecture Notes in Mathematics, 446 :5–49, 1975.

[BCC08] Adrien Blanchet, Vincent Calvez, and Jos´ A. Carrillo.
Convergence of the mass–transport steepest descent scheme for the subcritical Patlak–Keller–Segel model.
SIAM J. Numer. Anal., 46(2) :691–721, 2008.

[Ben] Sébastien Benzekry.
Mathematical analysis of a two-dimensional population model of metastatic growth including angiogenesis.
submitted, 2010.

[BGM09] Leonardo Baffico, Céline Grandmont, and Bertrand Maury.
Multiscale modelling of the respiratory tract.
Math. Models Methods Appl. Sci., 2009.

[Cho68] Alexandre J. Chorin.
Numerical solution of the Navier–Stokes equations.
Math. Comp., 22 :745–762, 1968.

[Cho69] Alexandre J. Chorin.
On the convergence of discrete approximations to the Navier–Stokes equations.
Math. Comp., 23 :341–353, 1969.

[CPSt] Vincent Calvez, Benoît Perthame, and Mohsen Sharifi tabar.
Modified Keller–Segel system and critical mass for the log interaction kernel.
Contemp. Math., 429 :45–62.

[DS09] Jean Dolbeault and Christian Schmeiser.
The two-dimensional Keller-Segel model after blow-up.
Discrete Contin. Dyn. Syst., 25(1) :109–121, 2009.

[FFS+05a] Maddalena Felici, Marcel Filoche, Christian Straus, Thomas Similowski, and Bernard Sapoval.
Diffusional screening in real 3D human acini–a theoretical study.
Respir Physiol Neurobiol, 145(2-3) :279–93, 2005.

[FFS05b] Marcel Filoche, Maddalena Felici, and Bernard Sapoval.
Renormalized random walk study of oxygen absorption in the human lung.
Physica A : Statistical Mechanics and its Applications, 357(1) :27, 2005.

[HGC+05] Chen Hou, Stefan Gheorghiu, Marc-Olivier Coppens, Virginia H. Huxley, and Peter Pfeifer.
Fractals in biology and medicine.
In *Gas Diffusion through the Fractal Landscape of the Lung : How Deep Does Oxygen Enter the Alveolar System ?*, pages 17–30, 2005.

[HLHOP] Frédéric Hecht, Antoine Le Hyaric, Koji Ohtsuka, and Olivier Pironneau.
Freefem++, finite elements software, `http ://www.freefem.org/ff++/`.

[IKS00] Kazuro Iwata, Kohkichi Kawasaki, and Nanako Shigesada.
A dynamical model for the growth and size distribution of multiple metastatic tumors.
Journal of Theoretical Biology, 203, Issue 2 :177–186, March 2000.

[Kho00] Boris N. Kholodenko.
Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades.
Eur J Biochem, 267(6) :1583–1588, Mar 2000.

[Mic05] Philippe Michel.
Principe d'entropie généralisée et dynamique de populations structurées.
Thèse de doctorat, Université Paris IX-Dauphine, november 2005.

[MnGNK09] Javier Muñoz García, Zoltan Neufeld, and Boris N. Kholodenko.
Positional information generated by spatially distributed signaling cascades.
PLoS Comput Biol, 5(3) :e1000330, 03 2009.

[Pai73] Manuel Paiva.
Gas transport in the human lung.
J Appl Physiol, 35(3) :401–410, 1973.

[Pou02] Frédéric Poupaud.
Diagonal defect measures, adhesion dynamics and Euler equation.
Methods Appl. Anal., 9(4) :533–561, 2002.

[RGBB09] Christophe Russo, Vasile Claudiu Giuraniuc, Ralf Blossey, and Jean-François Bodart.
On the equilibria of the mapk cascade : Cooperativity, modularity and bistability.
Physica A : Statistical Mechanics and its Applications, 388(24) :5070 – 5080, 2009.

[Rus09] Christophe Russo.
Dynamique et modularité de la voie Mitogen Activated Protein Kinase dans les ovocytes de Xénope : modélisation et approches expérimentales.
PhD thesis, Université de Lille 1, U.F.R. de Biologie, 2009.

[SFW02] Bernard Sapoval, Marcel Filoche, and Ewald R. Weibel.
Smaller is better–but not too small : a physical scale for the design of the mammalian pulmonary acinus.
Proc. Natl. Acad. Sci. U.S.A., 99(16) :10411–6, 2002.

[WSF05] Ewald R. Weibel, Bernard Sapoval, and Marcel Filoche.
Design of peripheral airways for efficient gas exchange.
Respiratory Physiology & Neurobiology, 148(1-2) :3 – 21, 2005.
Structure and Function in the Periphery of the Lung.

Première partie

Modèle pour une population de métastases

Chapitre 1

Un modèle décrivant la croissance et la distribution en taille d'une population de métastases

Le cancer est parmi les causes de mortalité les plus fréquentes dans les pays occidentaux. Une large part de la recherche médicale est donc logiquement dédiée à l'étude de cette maladie et la modélisation mathématique doit être considérée comme un outil supplémentaire au service des médecins et biologistes pour mieux comprendre les mécanismes du cancer et déterminer les traitements les mieux adaptés. Dans ce travail, fruit d'une collaboration avec THIERRY GOUDON et PAULINE LAFITTE, nous nous intéressons à un modèle déjà proposé en 2000, par IWATA *et al.* [IKS00] qui décrit l'évolution d'une population de tumeurs métastatiques non–traitée. Nous donnons une analyse mathématique du modèle qui nous ammène à la détermination d'un taux Malthusien caractérisant la croissance exponentielle de la population. Nous réalisons également une analyse numérique du système aux dérivées partielles issu du modèle. L'article est paru dans Discrete and Continuous Dynamical Systems Series B sous le titre *A model describing the growth and the size distribution of multiple metastatic tumors.* [DGL09]

1.1 Introduction

In this paper we are concerned with a mathematical model describing the growth of tumors. The model has the form of a PDE which looks like a conservation law, endowed with a boundary condition of non local type. Precisely, tumors are characterized by their size $x \geq 1$ and we are concerned with the behavior of the size distribution $\rho(t,x)$:

$$\int_u^v \rho(t,x)\,\mathrm{d}x$$

gives the number of tumors with size ranging in the domain $[u,v]$ at time $t \geq 0$. The dynamics are governed by the combination of two phenomena:
- a tumor of size x grows with a rate $w(x) \geq 0$,
- the growing tumor also emits new malignant cells with a rate $\beta(x) \geq 0$.

We further do not consider any treatment. Then, we are led to the following equation

$$\partial_t \rho + \partial_x(w\rho) = 0, \tag{1.1}$$

$$w(1)\rho(t,1) = \int_1^\infty \beta(y)\rho(t,y)\,\mathrm{d}y + \beta(x_p(t)). \tag{1.2}$$

In (1.2), the last term accounts for the contribution of the primary tumor whose evolution follows the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}x_p(t) = w(x_p(t)),\ x_p(0) = 1. \tag{1.3}$$

The problem is completed by assuming that initially there is no metastatic tumor

$$\rho_{|t=0} = 0. \tag{1.4}$$

The problem (1.1)-(1.4) appears exactly in that way in [IKS00]. The model is intended to describe the earlier stages of the disease when the number of metastases remains far below measurement capabilities. Therefore the crucial point is to be able to characterize the large time asymptotics of the size distribution: we shall see that the main information is embodied into a positive parameter, that will be denoted λ_0, characterizing the speed of spread of the cancer at this stage. In fact we show that

$$\rho(t,x) \sim C\, e^{\lambda_0 t} N(x),$$

as t tends to infinity. Having a sharp estimate of this parameter, which depends on the functions w and β, is therefore particularly important to prevent tumour invasion and to devise successful treatment strategies. In addition to the introduction of the model, in [IKS00], the solutions predicted by (1.1)-(1.4) are compared with size distributions successively measured by computed tomography images of a patient with hepatocellular carcinoma (liver cancer), showing the ability of the model to predict the tumors invasion.

System (1.1)-(1.4) belongs to the family of the so–called McKendrick-Von Foerster equations. Here, the assumptions on the coefficients are the following:

$$\begin{cases} w \in C^1(\mathbb{R}), \\ |w(x) - w(y)| \le L|x-y|, \\ w(x) > 0 \text{ for any } x \in [1,b[,\ w(b) = 0, \end{cases} \tag{1.5}$$

and for the emission rate we have

$$\beta \in L^\infty([1,b]),\ \beta(x) > 0 \text{ on } [1,b] \tag{1.6}$$

The parameter $1 < b < \infty$ represents the maximal size of a tumor: intuitively the larger the tumor, the more difficult it is to grow. In particular, the growth rate w vanishes at size b. As a consequence we note that

$$\int_1^b \frac{\mathrm{d}x}{w(x)} = \infty. \tag{1.7}$$

Compared to classical McKendrick-Von Foerster models, where typically $w = 1$, this singularity induces non trivial difficulties in the mathematical analysis.

A typical model is given by the Gompertzian law

$$w(x) = ax\ \ln(b/x), \tag{1.8}$$

for some $a > 0$ and then (1.3) can be integrated as follow:

$$x_p(t) = b^{1-e^{-at}}. \tag{1.9}$$

The Gompertz law (1825) is an empirical one, but there exist biological models of the tumor growth which explain characteristic Gompertz-type growth curves. See [FM86, GW89, GW90, KB03]. They are two-compartment models: they distinguish the proliferating cells and the quiescent ones. The behavior of these two types of cells and so the inter-compartment transition rate function, depends on the tumor size. In [GW89], it is shown that in the case of a tumor forming a necrotic center the model predicts that the tumor grows monotonically to its ultimate size according to a typical S-shaped Gompertz' curve. We should specify that in those models the cells death is already taken in account.

The expression of the emission rate β depends on the distribution of blood to the tumors: current models are $\beta(x) = m\,x^\alpha$. When the tumor is vascular superficially, that is the blood vessels distribute on its surface, α is

assigned to $2/3$ because the surface area of the tumor is proportional to $x^{2/3}$. On the other hand, when the blood vessel distribution is homogeneous in the tumor, α is assigned to 1. In this paper, α is dealt with as a parameter and for numerical simulations we adopt for α the value obtained in [IKS00] by comparison to clinical data ($\alpha = 0.663$).

The mathematical questions we address here can be summarized as follows:

- the definition of the parameter $\lambda_0 > 0$, which in the population dynamics context would be interpreted as a Malthusian rate, relies on the resolution of an eigenvalue problem, associated to Equations (1.1)-(1.2). Usually exhibiting a solution to this problem combines the positivity and compactness properties of the underlying operator, through an application of the Krein-Rutman theorem [Per07]. However, the fact that w vanishes at some point introduces some singularity (see e.g. (1.7)), and leads to technical difficulties in applying such a method. Hence, we shall use a more direct approach, thanks to a tricky change of unknowns which simplifies the question. We obtain then the asymptotic behavior of the solution by an entropy method, see [MMP04].

- the numerical simulation of the problem also presents some interesting difficulties since

  - the crucial question relies on a sharp evaluation of the asymptotic trend for large times and it thus requires some care in the numerical scheme we use;

  - the problem involves physical parameters ranging in an exponentially wide domain: in particular the typical value of b is very large (10^{11}) which makes the use of reasonably refined meshes non affordable.

The paper is organized as follows. At first, we discuss some properties of the PDE (1.1)-(1.4). In particular, we establish the well-posedness of the problem under relevant conditions. We also show that the problem can be reinterpreted as a standard initial boundary value problem, without source term in the boundary condition. Secondly, we detail the corresponding eigenvalue problem. In turn, we discuss the large time asymptotics by using entropy methods, as presented in [Mic05] and investigate the convergence rate. Finally, we describe how the problem can be handled numerically, so that we can compare on numerics the solution with the expected profile. It is worth mentioning that a similar program is addressed by Barbolosi, Benabdallah, Hubert and Verga [BBHV09] who use a completely different mathematical toolbox, both on the theoretical level where their proofs rely on semi-group arguments which yield very sharp results and on the numerical level since they use a different scheme based on characteristics (see also [ALM99]). We add an Appendix with a discrete approach of the problem and show that (1.1)-(1.4) can be derived from a semi-discrete problem.

1.2 Analysis of the PDE

1.2.1 Existence and uniqueness of a solution

Let us start by considering the evolution problem

$$\begin{cases} \partial_t f + \partial_x(wf) = 0, \\ f_{|t=0} = f_0, \end{cases} \tag{1.10}$$

where w fulfills (1.5). Since w has a positive value at $x = 1$, the problem (1.10) should be augmented with an incoming boundary condition

$$(wf)(t,1) = k(t) \geq 0. \tag{1.11}$$

We aim at solving the initial-boundary-value problem (1.10)-(1.11).

To this end, we introduce the characteristics associated to w, that is the solutions of the autonomous ODE:

$$\frac{\mathrm{d}}{\mathrm{d}t} X(t,x) = w(X(t,x)),\ X(0,x) = x. \tag{1.12}$$

We also set

$$J(-t,x) = \exp\left(-\int_{-t}^{0} w'(X(\sigma,x))\,\mathrm{d}\sigma\right) = \exp\left(\int_{0}^{t} w'(X(-\sigma,x))\,\mathrm{d}\sigma\right), \tag{1.13}$$

which is the jacobian of the change of variable $y = X(-t,x)$, i.e., $\mathrm{d}y = \partial_x X(-t,x)\,\mathrm{d}x = J(-t,x)\,\mathrm{d}x$. Finally, we define the exit time function

$$T_\star(x) = \inf\{t \geq 0, \text{ for any } 0 \leq s \leq t,\, X(-s,x) \in (1,b)\}.$$

In view of (1.5), we have $X(-T_\star(x),x) = 1$ when $T_\star(x)$ is finite. As a matter of fact, we note that

$$T_\star(x) = \int_0^{T_\star(x)} \frac{\frac{\mathrm{d}X}{\mathrm{d}s}(-s,x)}{w(X(-s,x))}\,\mathrm{d}s = \int_1^x \frac{\mathrm{d}y}{w(y)},$$

which tends to $+\infty$ as x tends to b owing to (1.7).

We integrate (1.10) along the characteristics, so that we get

$$\frac{d}{ds}\left[\ln\left(f(t+s,X(s,x))\right)\right] = -w'(X(s,x)).$$

Finally, we obtain the formula

$$f(t,x) = f_0(X(-t,x))\,J(-t,x)\,\mathbb{1}_{0\leq t\leq T_\star(x)} + k(t-T_\star(x))\,J(-T_\star(x),x)\mathbb{1}_{t\geq T_\star(x)}.$$

For given smooth data $f_0 \in C^1(\mathbb{R}^+)$, $k \in C^1(\mathbb{R}^+)$, this formula clearly defines the solution of (1.10)-(1.11). In particular, considering data f_0 supported in $[1,b]$, the support of the solution remains in $[1,b]$. Furthermore, it also makes sense in a more general context allowing discontinuous data or even measure valued solutions, through the dual formulation

$$\begin{aligned}\int_1^b \varphi(x) f(t,\mathrm{d}x) &= \int_1^b \varphi(X(t,y)) f_0(\mathrm{d}y) \\ &\quad + \int_1^b \varphi(x) k(t-T_\star(x))\,J(-T_\star(x),x)\,\mathbb{1}_{t\geq T_\star(x)}\,\mathrm{d}x,\end{aligned}$$

which holds for any $\varphi \in C^1([1,b])$. The last integral can be rewritten by using the change of variables

$$s = T_\star(x), \qquad x = X(s,1), \qquad \mathrm{d}x = w(X(s,1))\,\mathrm{d}s.$$

We end up with

$$\int_1^b \varphi(x) f(t,\mathrm{d}x) = \int_1^b \varphi(X(t,y)) f_0(\mathrm{d}y) + \int_1^t k(t-s)\,w\varphi(X(s,1))J^{-1}(s,1)\;\mathrm{d}s. \tag{1.14}$$

Definition 1.2.1 *Let $\mathcal{M}_+^1$ the positive cone of bounded measures on $[1,b]$, that is the set of continuous and non negative linear forms on $C^0([1,b])$.*

For given $f_0 \in \mathcal{M}_+^1$ and $k \in L^p(0,T)$ the formula (1.14) defines the unique solution $f \in C^0(0,T;\mathcal{M}^1-\text{weak}-\star)$ of (1.10)-(1.11). It means that for any $\varphi \in C^0([1,b])$, we have

- there exists $C_T > 0$ such that for any $0 \leq t \leq T$, $\left|\int_1^b \varphi(x) f(t,\mathrm{d}x)\right| \leq C_T\|\varphi\|_\infty$,
- the function $t \mapsto \int_1^b \varphi(x) f(t,\mathrm{d}x)$ is continuous on $[0,T]$.

We are thus led to the following definition.

Definition 1.2.2 *A measure-valued function $t \mapsto \rho(t,\mathrm{d}x)$ such that*

$$\rho \in L^\infty(\mathbb{R}^+;\mathcal{M}_+^1([1,b])) \cap C^0(0,T;\mathcal{M}^1 - \textit{weak} - \star)$$

is said a solution of

$$\begin{aligned}&\partial_t\rho + \partial_x(w\rho) = 0,\\ &w\rho(t,1) = \int_1^b \beta(y)\rho(t,\mathrm{d}y),\\ &\rho\big|_{t=0} = f_0,\end{aligned} \tag{1.15}$$

if (1.14) *is satisfied for any $\varphi \in C^0([1,b])$ with $k(t) = \int_1^b \beta(y)\rho(t,\mathrm{d}y) \in C^0(\mathbb{R}^+)$.*

Then, we use a standard fixed point procedure to find the solution of (1.10)-(1.11) with the self-consistent boundary condition

$$k(t) = \int_1^b \beta(y) f(t, dy).$$

We set the sequence $\left(f^{(n)}\right)_{n\in\mathbb{N}}$ defined by:

$$\begin{aligned} &\partial_t f^{(n+1)} + \partial_x (w f^{(n+1)}) = 0, \\ &w f^{(n+1)}(t,1) = \int_1^b \beta(y) f^{(n)}(t,\, \mathrm{d}y), \\ &f^{(n+1)}_{|t=0} = f_0, \end{aligned}$$

and as initial guess we choose

$$f^{(0)} \equiv 0.$$

We show that the sequence is non-decreasing and bounded. By induction the sequence $\left(f^{(n)}\right)_{n\in\mathbb{N}}$ is non-decreasing; in fact, for any $\varphi \in C^1([1,b])$, $\varphi \geq 0$,

$$\int_1^b \varphi(x)(f^{(1)} - f^{(0)})(t,\, \mathrm{d}x) = \int_1^b \varphi(X(t,y)) f_0(\,\mathrm{d}y) \geq 0,$$

and

$$\begin{aligned} &\int_1^b \varphi(x)(f^{(n+1)} - f^{(n)})(t,\, \mathrm{d}x) \\ &= \int_1^t \mathrm{d}s \left(\int_1^b \beta(y)(f^{(n)} - f^{(n-1)})(t-s,\, \mathrm{d}y) \right) w\varphi(X(s,0))\, J^{-1}(s,0). \end{aligned}$$

So we deduce that $f^{(n+1)} - f^{(n)} \geq 0$ for any $n \in \mathbb{N}$.

We prove the boundeness also by induction; assume that

$$\forall \varphi \in C^1([1,b]),\ \|\varphi\|_\infty \leq 1,\ \sup_{0\leq t<\infty} \left| \int_1^b \varphi(y) e^{-\mu t} f^{(n)}(t,\, \mathrm{d}y) \right| \leq M,$$

where $\mu > 0$ and $M > 0$ will be set further on. Then, we have for any test function φ with $\|\varphi\|_\infty \leq 1$,

$$\begin{aligned} &\left| \int_1^b \varphi(x) f^{(n+1)}(t,\, \mathrm{d}x) e^{-\mu t} \right| \\ &\leq \int_1^b |f_0(\,\mathrm{d}y)| e^{-\mu t} \\ &\quad + \left| \int_0^t \mathrm{d}s \left(\int_1^b \beta(y) e^{-\mu(t-s)} f^{(n)}(t-s,\, \mathrm{d}y) \right) w\varphi(X(s,0)) J^{-1}(s,0) e^{-\mu s} \right| \\ &\leq \int_1^b |f_0(\,\mathrm{d}y)| e^{-\mu t} + \int_0^t \|\beta\|_\infty M \|w\|_\infty\, J^{-1}(s,0) e^{-\mu s}\, \mathrm{d}s. \end{aligned} \tag{1.16}$$

Since, according to (1.5)

$$0 \leq J^{-1}(s,0) = \exp\left(\int_s^0 w'(X(\sigma,x))\, \mathrm{d}\sigma \right) \leq \exp(sL),$$

we have:

$$\left| \int_1^b \varphi(x) f^{(n+1)}(t,\, \mathrm{d}x) e^{-\mu t} \right| \leq \int_1^b |f_0(\,\mathrm{d}y)| e^{-\mu t} + \|\beta\|_\infty \|w\|_\infty \frac{M}{\mu - L}.$$

We choose μ and M such that

$$\alpha := \frac{\|\beta\|_\infty \|w\|_\infty}{\mu - L} < 1 \qquad \text{and} \qquad \frac{\int_1^b |f_0(\,\mathrm{d}y)|}{1-\alpha} \leq M.$$

Then

$$\sup_{0\le t<\infty}\left|\int_1^b \varphi(x) f^{(n+1)}(t,\,\mathrm{d}x)e^{-\mu t}\right| \le M,$$

holds for any $\varphi \in C^1([1,b])$ verifying $\|\varphi\|_\infty \le 1$. Therefore, for any $n \in \mathbb{N}$, we deduce that

$$\sup_{0\le t<\infty} e^{-\mu t}\|f^{(n)}(t)\|_{\mathcal{M}^1} \le M.$$

We end the proof by establishing a contraction property. We fix $\varphi = \beta$, and we get

$$\begin{aligned}
0 &\le \int_1^b \beta(x)(f^{(n+1)} - f^{(n)})(t,\,\mathrm{d}x)e^{-\mu t}\\
&= \int_0^t \mathrm{d}s\left(\int_1^b \beta(x)e^{-\mu(t-s)}(f^{(n)} - f^{(n-1)})(t-s,\,\mathrm{d}y)\right) w\beta(X(s,0))e^{-\mu s}\,J^{-1}(s,0)\\
&\le \sup_{\tau\ge 0}\left(e^{-\mu\tau}\int_1^b \beta(y)(f^{(n)} - f^{(n-1)})(\tau,\,\mathrm{d}y)\right)\int_0^t w\beta(X(s,0))e^{-\mu s}\,J^{-1}(s,0)\,\mathrm{d}s\\
&\le \sup_{\tau\ge 0}\left(e^{-\mu\tau}\int_1^b \beta(y)(f^{(n)} - f^{(n-1)})(\tau,\,\mathrm{d}y)\right)\frac{\|w\|_\infty\|\beta\|_\infty}{\mu - L}\\
&\le \alpha\,\sup_{\tau\ge 0}\left(e^{-\mu\tau}\int_1^b \beta(y)(f^{(n)} - f^{(n-1)})(\tau,\,\mathrm{d}y)\right),
\end{aligned}$$

where, by definition, $0 < \alpha < 1$. Hence, the sequence $\left(f^{(n)}\right)_{n\in\mathbb{N}}$ converges and the limit fulfills the requirements of Definition 1.2.2.

Theorem 1.2.3 *Assume* (1.5)-(1.6). *Let* $f_0 \in \mathcal{M}^1_+$. *Then, there exists a unique solution of* (1.15) *in the sense of Definition 1.2.2.*

Remark 1 *The proof also shows that if the initial data is absolutely continuous with respect to the Lebesgue measure, then the solution is absolutely continuous with respect to the Lebesgue measure too.*

At first look, this statement does not fit the resolution of the original system (1.1)-(1.4). However, we note that $\delta(x = X(t,0))$ satisfies (1.14) with $f_0 = \delta(x = 1)$ and $k = 0$. Accordingly we show that ρ satisfies (1.10)-(1.11) with $k(t) = \int_1^b \beta\rho(t,\,\mathrm{d}x) + \beta(X(t,0))$ and $f_0 = 0$ iff $f = \rho + \delta(x = X(t,0))$ satisfies (1.10)-(1.11) with $k(t) = \int_1^b \beta\rho(t,\,\mathrm{d}x)$ and $f_0 = \delta(x = 1)$. This remark shows that (1.1)-(1.4) is equivalent to (1.15) with initial data $\delta(x = 1)$. Note however that the interpretation is slightly changed since ρ represents secondary tumors only, while f represents the primary and secondary tumors.

Corollary 1 *Assume* (1.5)-(1.6). *Then, there exists a unique solution in the set* $L^\infty(\mathbb{R}^+;\mathcal{M}^1_+([1,b]))\cap C^0(\mathbb{R}^+;\mathcal{M}^1_+ - \mathit{weak} - \star))$ *to* (1.1)-(1.4).

1.2.2 A generational point of view

The total population of metastasis is composed of the daughters, the granddaughters, the great-granddaughters, etc. of the primary tumor, hence the total population of metastasis can be structured by the rank in the line of descent from the primary tumor and the total density of population can be seen as the sum of the density of each of these sub-populations. Let us denote by ρ_n the density of the population of rank n in the line of descent. Then ρ_1 will denote the population of the daughter-metastasis of the primary tumor. We have the following recursion:

$$\partial_t\rho_{n+1} + \partial_x(w\rho_{n+1}) = 0, \tag{1.17a}$$

$$\rho_{n+1}\big|_{t=0} = 0, \tag{1.17b}$$

$$w(1)\rho_{n+1}(t,1) = \int_1^b \beta(x)\rho_n(t,x)\,\mathrm{d}x, \tag{1.17c}$$

with the initialization:

$$\begin{aligned} \partial_t \rho_1 + \partial_x(w\rho_1) &= 0, \\ \rho_1\big|_{t=0} &= 0, \\ w(1)\rho_1(t,1) &= \beta(x_p(t)), \end{aligned}$$

where x_p is defined in (1.3). We set then

$$\rho = \sum_{n=1}^{\infty} \rho_n. \tag{1.18}$$

A proof very similar to the previous one, justifies the convergence of the series and shows that ρ defined in (1.18) is solution of (1.1)-(1.4). We refer to the fifth section for numerical simulations and a discussion on the contribution of each generation in the total population density.

1.3 Eigenproblem

In this section we are concerned with the eigenproblem associated to (1.15). We aim at establishing the existence of a unique positive eigenvalue associated to a positive eigenvector and dual eigenvector, namely we seek (λ_0, N, Φ) satisfying

$$\frac{\partial}{\partial x}\left(N(x)w(x)\right) + \lambda_0 N(x) = 0, \tag{1.19a}$$

$$w(1)N(1) = \int_1^b \beta(y)N(y)\,\mathrm{d}y, \tag{1.19b}$$

$$-w(x)\frac{\partial}{\partial x}\Phi(x) + \lambda_0\Phi(x) = \Phi(1)\beta(x), \tag{1.19c}$$

$$\lambda_0 > 0, \quad N \geq 0, \quad \Phi \geq 0, \quad \int_1^b N\Phi\,\mathrm{d}x = 1, \quad \int_1^b N\,\mathrm{d}x = 1. \tag{1.19d}$$

Theorem 1.3.1 *There exists a unique* $\lambda_0 > 0$ *and* $(N, \Phi) \in L^1([1,b]) \times L^\infty([1,b])$ *satisfying* (1.19a)-(1.19d).

Proof. We start by studying (1.19a)-(1.19b). We set the change of variable: $x = x_p(t)$, where x_p is defined by (1.3), and we define

$$U(t) = (wN)(x_p(t)).$$

Then (1.19a)-(1.19b) recasts as

$$\begin{cases} \partial_t U + \lambda_0 U = 0, \\ \\ U(0) = \displaystyle\int_0^\infty B(t)U(t)\,\mathrm{d}t, \end{cases} \tag{1.20}$$

where $B(t) = \beta(x_p(t))$. This mere ODE can be integrated and we get

$$U(t) = e^{-\lambda_0 t}\int_0^\infty B(s)U(s)\,\mathrm{d}s. \tag{1.21}$$

We have $\int_0^\infty BU\,\mathrm{d}s \neq 0$: indeed if $\int_0^\infty BU\,\mathrm{d}s = 0$, then, due to the non degeneracy condition (1.6), $U \equiv 0$ and U can not be a eigenvector. We multiply (1.21) by B and integrate:

$$\int_0^\infty B(t)U(t)\,\mathrm{d}t = \int_0^\infty B(t)e^{-\lambda_0 t}\,\mathrm{d}t \cdot \int_0^\infty B(s)U(s)\,\mathrm{d}s.$$

Since $\int_0^\infty BU\,\mathrm{d}s \neq 0$, we deduce that

$$\int_0^\infty B(t)e^{-\lambda_0 t}\,\mathrm{d}t = 1.$$

We set

$$F(\mu) := \int_0^\infty B(t)e^{-\mu t}\,\mathrm{d}t.$$

Clearly, F is decreasing and we note that $F(\infty) = 0$ since B is bounded, and

$$F(0) = \int_0^\infty B(t)\,\mathrm{d}t = \int_1^b \frac{\beta(x)}{w(x)}\,\mathrm{d}x = +\infty,$$

(see (1.7)). We conclude that

$$\exists!\ \lambda_0 \in (0,\infty) \text{ s.t. } F(\lambda_0) = 1, \tag{1.22}$$

and, according to (1.21), $U(t) = U(0)e^{-\lambda_0 t}$ is a solution of (1.20) associated to the eigenvalue λ_0.

Next, we check that the eigenspace associated to λ_0 is of dimension one. Let $v(t)$ be another eigenvector associated to λ_0. Since we can write

$$v(t) = CU(t) + \bar{v}(t),$$

where $\bar{v}$ is such that $\int_0^\infty \bar{v}(t)U(t)\,\mathrm{d}t = 0$, we can assume that $\int_0^\infty e^{-\lambda_0 s}v(s)\,\mathrm{d}s = 0$. Moreover, we seek v in $W^{1,1}(\mathbb{R}^+)$ and assume $v \neq 0$. Then, we have

$$\begin{aligned}\int_0^\infty v'(s)\,\mathrm{d}s &= -v(0)\\ &= \left[e^{-\lambda_0 t}v(t)\right]_0^\infty = \int_0^\infty \Big(-\lambda_0 e^{-\lambda_0 t}v(t) + e^{-\lambda_0 t}v'(t)\Big)\,\mathrm{d}t\\ &= -2\lambda_0\int_0^\infty e^{-\lambda_0 t}v(t)\,\mathrm{d}t = 0.\end{aligned}$$

However, (1.20) implies that

$$\int_0^\infty v'(s)\,\mathrm{d}s = -v(0) = -\int_0^\infty B(s)v(s)\,\mathrm{d}s,$$

which yields a contradiction since $\int_0^\infty B(s)v(s)\,\mathrm{d}s \neq 0$ otherwise v would be identically 0 by (1.21). We conclude that the eigenspace associated to λ_0 is $\mathrm{Span}\{t \mapsto e^{-\lambda_0 t}\}$. Coming back to the original variables yields

$$N(x) = \frac{U(0)}{a\ln b}\,\frac{1}{x}\left(1 - \frac{\ln x}{\ln b}\right)^{\lambda_0/a-1}, \tag{1.23}$$

which is integrable since $\frac{\lambda_0}{a} - 1 > -1$ (but the behavior for x next to b depends on the sign of $\frac{\lambda_0}{a} - 1$).

Turning now to the dual problem (1.19c), we set again the change of variable $x = x_p(t)$, and we get for $\Psi(t) = \Phi(x_p(t))$,

$$-\frac{\partial\Psi}{\partial t} + \lambda_0\Psi = \Psi(0)B(t),$$

where λ is defined by (1.22), which means that:

$$\Psi(t) = e^{\lambda_0 t}\,\Psi(0)\left[1 - \int_0^t B(u)e^{-\lambda_0 u}du\right].$$

Since, according to (1.22), $\int_0^\infty B(u)e^{-\lambda_0 u}du = 1$ and $B(u)e^{-\lambda_0 u} \geq 0$, we have $\Psi > 0$. Moreover,

$$\lim_{t\to\infty}\Psi(t) = 0.$$

In fact,

$$\Psi(t) = \Psi(0)\int_t^\infty B(u)e^{-\lambda_0(u-t)}\,\mathrm{d}u \leq \Psi(0)\int_t^\infty B(u)\,\mathrm{d}u.$$

and since $B \in L^1([0,\infty))$, the right hand side tends to 0 as t tends to infinity. This shows that Ψ is bounded, so $\Phi \in L^\infty([1,b])$. To satisfy the normalization conditions, we choose convenient constants $U(0)$ and $\Psi(0)$.

□

1.4 General relative entropy and asymptotic behavior

The general relative entropy method is currently used for McKendrick-Von Foerster equations (see [Mic05]) to find *a priori* bounds and to describe the asymptotic behavior of the solution. It is based on the eigenproblem described before.

1.4.1 Conservation law and general relative entropy

Theorem 1.4.1 *Let* (N, λ_0, Φ) *be a solution of the problem* (1.19a)-(1.19d)*, then for any* $\rho(0,\cdot) \geq 0$ *and any* $\rho(t,\cdot)$ *solution of* (1.15) *with initial data* $\rho(0,\cdot) \in L^\infty\left([1,b], \frac{\mathrm{d}x}{N}\right)$*,*

1. *we have the conservation law*

$$\int_1^b \Phi(x)\rho(t,x)e^{-\lambda_0 t}\,\mathrm{d}x = \int_1^b \Phi(x)\rho(0,x)\,\mathrm{d}x, \tag{1.24}$$

2. *for any convex function* H*, the general relative entropy defined as*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_1^b \Phi(x)N(x)H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}x := -D_H, \tag{1.25}$$

with

$$\begin{aligned} -D_H &= w(1)N(1)\Phi(1) \\ &\quad\times\left[H\left(\int_1^b \frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\,\mathrm{d}\mu(x)\right) - \int_1^b H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}\mu(x)\right], \end{aligned} \tag{1.26}$$

where

$$\mathrm{d}\mu(x) = \frac{\beta(x)N(x)}{w(1)N(1)}\,\mathrm{d}x,$$

is non positive.

Proof. Remark first that the conservation law can be obtained from the entropy formula, by choosing $H(x) = x$, since $-D_H$ is zero in this case. In order to obtain (1.25) we develop its left hand side:

$$\begin{aligned} -D_H &= \frac{\mathrm{d}}{\mathrm{d}t}\int_1^b \Phi(x)N(x)H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}x, \\ &= \int_1^b \Phi N\left[\partial_t\rho\frac{e^{-\lambda_0 t}}{N} - \lambda N\frac{\rho e^{-\lambda_0 t}}{N^2}\right]H'\left(\frac{\rho e^{-\lambda_0 t}}{N}\right)\mathrm{d}x, \\ &= \int_1^b \Phi N\left[-\partial_x(w\rho)\frac{e^{-\lambda_0 t}}{N} + \partial_x(wN)\frac{\rho e^{-\lambda_0 t}}{N^2}\right]H'\left(\frac{\rho e^{-\lambda_0 t}}{N}\right)\mathrm{d}x, \\ &= \int_1^b \Phi N\left[-w(\partial_x\rho)\frac{e^{\lambda_0 t}}{N} + (\partial_x N)w\frac{\rho e^{-\lambda_0 t}}{N^2}\right]H'\left(\frac{\rho e^{-\lambda_0 t}}{N}\right)\mathrm{d}x, \\ &= \int_1^b \Phi N w\,\partial_x\left(\frac{\rho e^{-\lambda_0 t}}{N}\right)H'\left(\frac{\rho e^{\lambda_0 t}}{N}\right)\mathrm{d}x, \\ &= w(1)\Phi(1)N(1)H\left(\frac{\rho(t,1)e^{-\lambda_0 t}}{N(1)}\right) + \int_1^b \Phi(1)\beta N\,H\left(\frac{\rho e^{-\lambda_0 t}}{N}\right)\mathrm{d}x, \\ &= w(1)N(1)\Phi(1)\left[H\left(\int_1^b \frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\,\mathrm{d}\mu(x)\right)\right. \\ &\qquad\left. - \int_1^b H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}\mu(x)\right]. \end{aligned}$$

Therefore D_H is non negative by virtue of the Jensen Lemma since H is convex and $\mathrm{d}\mu$ is a probability measure.

□

Proposition 1 *For $H(x) = x^2$, D_H vanishes iff $\frac{\rho(t,\cdot)e^{-\lambda_0 t}}{N(\cdot)} = C(t)$ does not depend on x.*

Proof. It is a consequence of the equality case in the Cauchy–Schwarz inequality.

□

1.4.2 Boundedness properties

We deduce from the fact that entropy is decreasing some *a priori* bounds, according to [MMP04], [Mic05]. In fact, by using Theorem 1.4.1, we get:

• L^∞ **bound.** Let $C \geq 0$. Then, choosing $H(x) = (x - C)^+$ which is convex non negative, we get

$$\rho(0,\cdot) \leq CN(\cdot) \implies \forall t \geq 0, \quad \rho(t,\cdot)e^{-\lambda_0 t} \leq CN(\cdot). \tag{1.27}$$

• L^∞ **bound.** Let $c \geq 0$. Choosing $H(x) = (c - x)^+$ which is convex non negative, we get in the same way

$$\rho(0,\cdot) \geq cN(\cdot) \implies \forall t \geq 0, \quad \rho(t,\cdot)e^{-\lambda_0 t} \geq cN(\cdot). \tag{1.28}$$

• L^p **bound.** Choosing $H(x) = |x|^p$ which is convex non negative, we finally get

$$\frac{\rho(0,\cdot)}{N(\cdot)} \in L^p([0,b], N\Phi\,\mathrm{d}x) \implies \forall t \geq 0, \quad \frac{\rho(t,\cdot)e^{-\lambda_0 t}}{N(\cdot)} \in L^p([0,b], N\Phi\,\mathrm{d}x) \tag{1.29}$$

and

$$\forall t \geq 0, \quad \int_1^b \left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)^p N(x)\Phi(x)\,\mathrm{d}x \leq \int_1^b \left(\frac{\rho(0,x)}{N(x)}\right)^p N(x)\Phi(x)\,\mathrm{d}x.$$

1.4.3 Asymptotic behavior

In this Section, we investigate the large time asymptotics. We describe first the asymptotic behaviour and then we detail the convergence rate to the asymptotic.

Long time asymptotic

Theorem 1.4.2 *For any $\rho(0,\cdot) \geq 0$ such that $\frac{\rho(0,\cdot)}{N(\cdot)} \in L^\infty \cap L^2([0,b], N\Phi\,\mathrm{d}x)$ and any $\rho(t,\cdot)$ solution of* (1.15) *with initial data $\rho(0,\cdot)$ then the following convergence*

$$\frac{\rho(\cdot,t)e^{-\lambda_0 t}}{N(\cdot)} \xrightarrow[t\to\infty]{} \int_1^b \Phi(x)\rho(0,x)\,\mathrm{d}x, \tag{1.30}$$

holds in $L^\infty \cap L^2([0,b], N\Phi\,\mathrm{d}x)$ as t tends to ∞.

The proof is based on the entropy dissipation, still following the approach of [MMP04]; for a similar result obtained by using semi-group techniques, we refer to [BBHV09].

Notation 1 *In this section, given $T \geq 0$,we will use the following notations:*

$$\begin{aligned} -D_{H,T}(t) &= \int_1^b H\left(\frac{\rho(t+T,x)e^{-\lambda_0(t+T)}}{N(x)}\right) \mathrm{d}\mu(x) \\ &\quad -H\left(\int_1^b \frac{\rho(t+T,x)e^{-\lambda_0(t+T)}}{N(x)}\,\mathrm{d}\mu(x)\right) \end{aligned}$$

and

$$v_T(t,x) = \rho(t+T,x)e^{-\lambda_0(t+T)}.$$

We start by proving the following lemma:

Lemma 1.4.3 *For any convex function H, we have*

$$\forall t \geq 0, \quad \lim_{T\to\infty} \int_0^t D_{H,T}(s)\,\mathrm{d}s = 0. \tag{1.31}$$

Proof. Writing (1.25)-(1.26) at time $t = s + T$ yields for any $T \geq 0$

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_1^b \Phi(x)N(x)H\left(\frac{\rho(s+T,x)e^{-\lambda_0(s+T)}}{N(x)}\right)\mathrm{d}x + w(1)\Phi(1)N(1)D_{H,T}(s) = 0.$$

By integration over $s \in (0,t)$, we get:

$$\begin{aligned} &\int_1^b \Phi(x)N(x)H\left(\frac{\rho(t+T,x)e^{-\lambda_0(t+T)}}{N(x)}\right)\mathrm{d}x + w(1)\Phi(1)N(1)\int_0^t D_{H,T}(s)\,\mathrm{d}s \\ &= \int_1^b \Phi(x)N(x)H\left(\frac{\rho(T,x)e^{-\lambda_0 T}}{N(x)}\right)\mathrm{d}x. \end{aligned} \tag{1.32}$$

Since

$$t \mapsto \int_1^b \Phi(x)N(x)H\left(\frac{\rho(t,x)e^{-\lambda_0 t}}{N(x)}\right)\mathrm{d}x$$

is non increasing and positive, it has a limit as t tends to infinity. Passing to the limit $T \to \infty$ in (1.32), we get (1.31).

□

In what follows, we fix $H(x) = x^2$. According to Lemma 1.4.3, we have:

$$\lim_{T\to\infty}\int_0^t\left[\int_1^b\left(\frac{v_T(s,x)}{N}\right)^2\mathrm{d}\mu(x) - \left(\int_1^b \frac{v_T(s,x)}{N}\,\mathrm{d}\mu(x)\right)^2\right]\mathrm{d}s = 0. \tag{1.33}$$

Next, we shall investigate the passage to the limit in both integrals arising in (1.33).

Lemma 1.4.4 *Assume that $\dfrac{\rho(0,\cdot)}{N(\cdot)} \in L^2([0,b], N\Phi\,\mathrm{d}x)$ then for any $t \geq 0$ it exists a sequence $(T_n)_n \to \infty$ and $v \in L^\infty([0,t], L^2\left([1,b], \frac{\Phi}{N}\right))$ such that*

$$\liminf_{T_n\to\infty}\int_0^t\int_1^b\left(\frac{v_{T_n}(s,x)}{N(x)}\right)^2\mathrm{d}\mu(x)\,\mathrm{d}s \geq \int_0^t\int_1^b\left(\frac{v(s,x)}{N(x)}\right)^2\mathrm{d}\mu(x)\,\mathrm{d}s.$$

Proof. According to (1.29), $\left(v_T(s,\cdot)\sqrt{\frac{\Phi(\cdot)}{N(\cdot)}}\right)$ is bounded in $L^2((0,t)\times(1,b))$, uniformly with respect to $T \geq 0$. Therefore, by the Banach-Alaoglu theorem, there exists a sequence $(T_n)_n$ and $v \in L^2\left((0,t)\times(1,b), \frac{\Phi(\cdot)}{N(\cdot)}\,\mathrm{d}x\,\mathrm{d}s\right)$ such that

$$v_{T_n}\sqrt{\frac{\Phi}{N}} \rightharpoonup v\sqrt{\frac{\Phi}{N}},$$

where the convergence holds weakly in $L^2((0,T)\times(1,b))$. Moreover, we have

$$\begin{aligned}\int_0^t\int_1^b v_{T_n}^2(s,x)\frac{\beta(x)}{N(x)}\,\mathrm{d}x\,\mathrm{d}s &= \int_0^t\int_1^b\left(v_{T_n}(s,x)\sqrt{\frac{\Phi(x)}{N(x)}}\right)^2\frac{\beta(x)}{\Phi(x)}\,\mathrm{d}x\,\mathrm{d}s,\\ &\geq \int_0^t\int_1^b\left(v_{T_n}(s,x)\sqrt{\frac{\Phi(x)}{N(x)}}\right)^2\Psi_\eta(x)\,\mathrm{d}x\,\mathrm{d}s,\end{aligned}$$

where we denote by Ψ_η the function

$$\Psi_\eta(x) := \frac{\beta(x)}{\Phi(x)}\mathbb{1}_{x\leq b-\eta}.$$

By convexity, for any $\varphi \in L^\infty([1,b])$, $\varphi \geq 0$, we have

$$\begin{aligned}&\liminf_{n\to\infty}\int_0^t\int_1^b\left(v_{T_n}(s,x)\sqrt{\frac{\Phi(x)}{N(x)}}\right)^2\varphi(x)\,\mathrm{d}x\,\mathrm{d}s\\ &\geq \int_0^t\int_1^b\left(v(s,x)\sqrt{\frac{\Phi(x)}{N(x)}}\right)^2\varphi(x)\,\mathrm{d}x\,\mathrm{d}s.\end{aligned}$$

In particular it holds for $\varphi = \Psi_\eta$ so that

$$\liminf_{n\to\infty}\int_0^t\int_1^b v_{T_n}^2(s,x)\frac{\beta(x)}{N(x)}\,\mathrm{d}x\,\mathrm{d}s \geq \int_0^t\int_1^b\left(v(s,x)\sqrt{\frac{\Phi(x)}{N(x)}}\right)^2\Psi_\eta(x)\,\mathrm{d}x\,\mathrm{d}s,$$

We pass to the limit in the right hand side, using the monotone convergence theorem as η tends to 0. We get

$$\liminf_{n\to\infty}\int_0^t\int_1^b v_{T_n}^2(s,x)\frac{\beta(x)}{N(x)}\,\mathrm{d}x\,\mathrm{d}s \geq \int_0^t\int_1^b\left(\frac{v(s,x)}{N(x)}\right)^2 N(x)\beta(x)\;\mathrm{d}x\,\mathrm{d}s,$$

which proves the lemma.

□

Lemma 1.4.5 *Assume that there exists a constant $C \geq 0$, such that $\rho(0,.) \leq CN(\cdot)$, then for any $t \geq 0$ there exists a sequence $(T_n)_n \to \infty$ and $v \in C^0([0,t], \mathcal{M}^1([1,b] - weak - *)$ such that*

$$\int_1^b \frac{v_{T_n}(s,x)}{N(x)}\,\mathrm{d}\mu(x) \longrightarrow \int_1^b \frac{v(s,x)}{N(x)}\,\mathrm{d}\mu(x),$$

uniformly on $[0,t]$ as $(T_n)_n$ tends to ∞.

Proof. For any $\varphi \in C^1([1,b])$, we set

$$V_T^{[\varphi]}(s) = \int_1^b v_T(s,x)\,\varphi(x)\,\mathrm{d}x.$$

Clearly, we have

$$\left|V_T^{[\varphi]}(s)\right| \le C\|\varphi\|_\infty, \tag{1.34}$$

owing to (1.27). Next, we compute

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}V_T^{[\varphi]}(s) &= \int_1^b \left[\partial_s(\rho(s+T,x))e^{-\lambda_0(s+T)}\varphi(x) - \lambda_0\rho(s+T,x)e^{-\lambda_0(s+t)}\varphi(x)\right]\mathrm{d}x,\\
&= \int_1^b w(x)\rho(S+T,x)e^{-\lambda_0(s+T)}\varphi'(x)\,\mathrm{d}x + w(1)\varphi(1)\rho(s+T,1)e^{-\lambda_0(s+T)}\\
&\quad - \lambda_0\int_1^b \rho(s+T,x)e^{-\lambda_0(s+T)}\varphi(x)\,\mathrm{d}x,\\
&= \int_1^b \left[w(x)\varphi'(x) - \lambda_0\varphi(x) + \varphi(1)\beta(x)\right]v_T(s,x)\,\mathrm{d}x.
\end{aligned}$$

Since $\|w\|_\infty = ab/e$, by using (1.27) again, we get,

$$\forall\varphi\in C^1([1,b]), \qquad \left|\frac{\mathrm{d}}{\mathrm{d}s}V_T^{[\varphi]}(s)\right| \le C\|\varphi\|_{W^{1,\infty}},$$

which proves that $\left(s\mapsto V_T^{[\varphi]}(s)\right)_{T\ge 0}$ is equicontinuous on $[0,t]$ for any $t>0$. Therefore, by the Arzela–Ascoli theorem, for any $\varphi\in C^1([1,b])$ it belongs to a compact set of $C^0([0,t])$. By (1.34) and a density argument, this is also true for any $\varphi\in C^0([1,b])$. Then using both separability and the diagonal Cantor process, we can find a subsequence T_n, and $v\in C^0([0,\infty),\mathcal{M}^1([1,b]-weak-*)$ such that the following convergence

$$V_{T_n}^{[\varphi]}(s) \longrightarrow \int_1^b v(s,x)\varphi(x)\,\mathrm{d}x,$$

as (T_n) tends to infinity, holds uniformly on $[0,t]$, for any $\varphi\in C^0([1,b])$ and any $t\ge 0$. Then we fix $\varphi = \beta/(N(1)w(1))$ and it proves the lemma.

□

Note that we can identify the limit v arising in Lemma 1.4.4 and 1.4.5 and Lemma 1.4.5 allows to pass to the limit in the second term in (1.33). Now, we go back to (1.33); using Lemma 1.4.4 and 1.4.5, we obtain

$$\begin{aligned}
0 &= \lim_{T_n\to\infty}\int_0^t\left[\int_1^b\left(\frac{v_{T_n}(s,x)}{N}\right)^2\mathrm{d}\mu(x) - \left(\int_1^b\frac{v_{T_n}(s,x)}{N}\,\mathrm{d}\mu(x)\right)^2\right]\mathrm{d}s,\\
&\ge \liminf_{T_n\to\infty}\int_0^t\int_1^b\left(\frac{v_{T_n}(s,x)}{N}\right)^2\mathrm{d}\mu(x)\,\mathrm{d}s - \lim_{T_n\to\infty}\int_0^t\left(\int_1^b\frac{v_{T_n}(s,x)}{N}\,\mathrm{d}\mu(x)\right)^2\mathrm{d}s,\\
&\ge \int_0^t\int_1^b\left(\frac{v(s,x)}{N}\right)^2\mathrm{d}\mu(x)\,\mathrm{d}s - \int_0^t\left(\int_1^b\frac{v(s,x)}{N}\,\mathrm{d}\mu(x)\right)^2\mathrm{d}s \ge 0.
\end{aligned}$$

We conclude by using Proposition 1 to obtain that the following convergence

$$\frac{\rho(\cdot,t)e^{-\lambda_0 t}}{N(\cdot)} \to C, \tag{1.35}$$

holds in $L^\infty\cap L^2([0,b],N\Phi\,\mathrm{d}x)$ as t tends to infinity. The constant is indeed uniquely determined by using the conservation law (1.24) and the normalization condition imposed on the eigenvectors (1.19d). In fact Proposition 1 shows that

$$\frac{v(s,\cdot)}{N(\cdot)} = C(s).$$

But (1.24) gives us

$$\int_1^b v_{T_n}(s,x)\Phi(x)\,\mathrm{d}x = \int_1^b \rho(0,x)\Phi(x)\,\mathrm{d}x,$$

and using arguments similar to the ones detailed in the proof of Lemma 1.4.5, we can pass to the limit and get

$$\int_1^b v(s,x)\Phi(x)\,\mathrm{d}x = \int_1^b \rho(0,x)\Phi(x)\,\mathrm{d}x,$$

which, using (1.19d), leads us to

$$\int_1^b N(x)\Phi(x)C(s)\,\mathrm{d}x = C(s) = \int_1^b \rho(0,x)\Phi(x)\,\mathrm{d}x.$$

Then $C(s)$ does not depend on s and is uniquely defined. We conclude that the whole sequence converge with

$$C = \int_1^b \Phi(x)\rho(0,x)\,\mathrm{d}x. \tag{1.36}$$

This completes the proof of Theorem 1.4.2.

Convergence rate

We are now interested in the rate of convergence of (1.30). The method we use is based on the invariant (1.24) and is inspired from [GP06].

In order to state our convergence result we need to rewrite the equation after renormalization by the exponential growth of the system. We define

$$\tilde{\rho}(t,x) = \rho(t,x)e^{-\lambda_0 t}.$$

Then, $\tilde{\rho}(t,x)$ satisfies the equations

$$\begin{cases} \partial_t\tilde{\rho}(t,x) + \partial_x(w(x)\tilde{\rho}(t,x)) + \lambda_0\tilde{\rho}(t,x) = 0, \\ w(1)\tilde{\rho}(t,0) = \displaystyle\int_1^b \beta(y)\tilde{\rho}(t,y)\,\mathrm{d}y, \end{cases}$$

with initial condition $\tilde{\rho}(t=0,x) = \tilde{\rho}^0(x) \in L^1([1,b])$. As a consequence of the general relative entropy principle we have the conservation law (see (1.24))

$$\frac{d}{dt}\int_1^b \tilde{\rho}(t,x)\Phi(x) = 0. \tag{1.37}$$

Finally, we introduce the function

$$h(t,x) = \tilde{\rho}(t,x) - \left(\int_1^b \rho^0(y)\Phi(y)\,\mathrm{d}y\right)N(x).$$

Then $h(t,x)$ fulfills

$$\begin{cases} \partial_t h(t,x) + \partial_x(w(x)h(t,x)) + \lambda_0 h = 0, \\ w(1)h(1) = \displaystyle\int_1^b \beta(x)h(t,x)\,\mathrm{d}x, \end{cases} \tag{1.38}$$

with initial condition $h(t=0,x) = h^0(x)$.

Note that the conservation law (1.37) implies that

$$\forall t \geq 0,\ \int_1^b h(t,x)\Phi(x)\,\mathrm{d}x = 0. \tag{1.39}$$

In what follows $X(t,x)$ designed the characteristic defined at (1.12) and $J(-t,x)$ is the jacobian of the change of variable $y = X(-t,x)$ (see (1.13)). The key assumption reads :

Assumption 1 *The following property holds*

$\exists \mu_0 > 0$, *and* $y \in [0, \infty)$, *such that,*$\forall u \in [1, b]$, $u \geq X(y, 1)$, *we have*

$$\Phi(1)\beta(u)N(u) \geq \mu_0 N(X(-y,u))\Phi(X(-y,u))J(-y,u). \tag{1.40}$$

We have the following result concerning the convergence rate. It is exponential.

Theorem 1.4.6 *The solutions to* (1.38) *satisfy*

$$\int_{X(y,1)}^{b} \frac{h(t,u)}{N(u)} N(X(-y,u))\Phi(X(-y,u))J(-y,u)\,\mathrm{d}u = 0,\ \forall t \geq y \geq 0. \tag{1.41}$$

Additionally, assume that Assumption 1 holds, then

$$\int_1^b |h(t,x)|\Phi(x)\,\mathrm{d}x \leq \min\left\{1, e^{-\mu_0(t-y)}\right\} \int_1^b |h^0(x)|\Phi(x)\,\mathrm{d}x.$$

Remark 2 *In our case it is easy to check that Assumption 1 is fulfilled with* $y = 0$ *and* $\mu_0 = m$, *but this choice can be improved choosing* $y \neq 0$. *In fact one can note that after the change of variable* $x = x_p(t)$ *already used at Section 1.3, the condition* (1.40) *is equivalent to*

$$e^{-\lambda_0(t+y)}B(t+y) \geq \mu_0 \int_t^\infty e^{-\lambda_0 u}B(u)\,\mathrm{d}u \tag{1.42}$$

where $B(t) = \beta(x_p(t))$. *We set:*

$$\Upsilon\ :\ t \in [0,\infty)\ \mapsto\ \int_t^\infty e^{-\lambda_0 u}B(u)\,\mathrm{d}u$$

and

$$f\ :\ t \in [0,\infty)\ \mapsto\ \frac{-\Upsilon'(t+y)}{\Upsilon(t)}.$$

A short analysis shows that $\Upsilon''(t)$ *is negative for* $0 \leq t < t_0$ *and positive if* $t > t_0$, *where*

$$t_0 = \frac{1}{a}\ln\left(\frac{a\alpha \ln b}{\lambda_0}\right).$$

For $y \geq t_0$, f *is then a non-decreasing function and for all* $t \geq 0$, $f(t) \geq f(0) = -\Upsilon'(y)$. *Thus, we can choose* $\mu_0 = -\Upsilon'(y)$ *which is maximal for* $y = t_0$ $(-\Upsilon'(t_0) > m)$.

Proof. The proof is an adaptation of the one proposed in [GP06], taking into account the characteristics due to the non constant speed rate w.

1st step. In exactly the same way as in [GP06], we show that

$$\frac{d}{dt}\int_1^b |h(t,x)|\Phi(x)\,\mathrm{d}x = \Phi(1)\left(\left|\int_1^b \beta(x)h(t,x)\,\mathrm{d}x\right| - \int_1^b \beta(x)|h(t,x)|\,\mathrm{d}x\right), \tag{1.43}$$

which provides the property:

$$t \mapsto \int_1^b |h(t,x)|\Phi(x)\,\mathrm{d}x \quad \text{is monotone nonincreasing.} \tag{1.44}$$

2nd step. We have

$$\frac{\partial}{\partial t}\frac{h(t,x)}{N(x)} + w(x)\frac{\partial}{\partial x}\frac{h(t,x)}{N(x)} = 0,$$

therefore we also have

$$\frac{h(t-y,x)}{N(x)} = \frac{h(t,X(y,x))}{N(X(y,x))}, \quad \text{for } t \geq y.$$

Thus, using (1.39), we get:

$$\begin{aligned} 0 &= \int_1^b \frac{h(t-y,x)}{N(x)}\Phi(x)N(x)\,\mathrm{d}x \\ &= \int_1^b \frac{h(t,X(y,x))}{N(X(y,x))}N(x)\Phi(x)\,\mathrm{d}x \\ &= \int_{X(y,1)}^b \frac{h(t,u)}{N(u)}N(X(-y,u))\Phi(X(-y,u))J(-y,u)\,\mathrm{d}u, \quad \forall t \geq y, \end{aligned}$$

and (1.41) is proved.

3rd step. Using (1.41) and Assumption 1, for all $t \geq y$, it holds

$$\begin{aligned} &\left|\int_1^b \Phi(1)\beta(x)h(t,x)\,\mathrm{d}x\right| \\ =&\left|\int_1^b \Phi(1)\beta(x)h(t,x)\,\mathrm{d}s \right. \\ &\qquad \left. -\mu_0\int_{X(y,1)}^b \frac{h(t,u)}{N(u)}N(X(-y,u))\Phi(X(-y,u))J(-y,u)\,\mathrm{d}u\right| \\ \leq &\int_1^b \left[\Phi(1)\beta(u)N(u)\right. \\ &\qquad \left. -\mu_0 \mathbb{1}_{u\geq X(y,1)}N(X(-y,u))\Phi(X(-y,u))J(-y,u)\right]\frac{|h(t,u)|}{N(u)}\,\mathrm{d}x. \end{aligned}$$

Combining this inequality and (1.43), we get

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\int_1^b |h(t,x)|\Phi(x)\,\mathrm{d}x \\ \leq &-\mu_0\int_{X(y,1)}^b \frac{|h(t,u)|}{N(u)}N(X(-y,u))\Phi(X(-y,u))J(-y,u)\,\mathrm{d}u \\ \leq &-\mu_0\int_1^b \frac{|h(t,X(y,x))|}{N(X(y,x))}N(x)\Phi(x)\,\mathrm{d}x \\ = &-\mu_0\int_1^b \frac{|h(t-y,x)|}{N(x)}N(x)\Phi(x)\,\mathrm{d}x \\ = &-\mu_0\int_1^b |h(t-y,x)|\Phi(x)\,\mathrm{d}x \\ \leq &-\mu_0\int_1^b |h(t,x)|\Phi(x)\,\mathrm{d}x, \quad \forall t \geq y, \end{aligned}$$

using the monotonicity property. Finally, a simple integration gives the convergence result. □

Remark 3 *In [IKS00], the exact solution of the problem is calculated explicitly using Laplace's transform and some complex analysis tools. The function is then given by the formula (See Appendix 1.8, page 61 for more details concerning the computation using Laplace's transform):*

$$\rho(x,t) = \frac{a}{mb^\alpha \ln b}\frac{1}{x}\sum_{k=0}^{\infty} e^{\lambda_k t}\left(1-\frac{\ln x}{\ln b}\right)^{\lambda_k/a-1}\frac{1}{c(\lambda_k)}, \tag{1.45}$$

$$\textit{for} \quad 1 \leq x < x_p(t),$$

where

$$c(\lambda_k) = \sum_{n=0}^{\infty} \frac{(-\alpha \ln b)^n}{n!(\lambda_k/a + n)^2}, \tag{1.46}$$

and the λ_k's are the roots of:

$$\frac{1}{m}\lambda_k = \sum_{n=0}^{\infty} \frac{(\alpha \ln b)^n}{(\lambda_k/a + 1)\dots(\lambda_k/a + n)} = \mathcal{F}(\lambda_k, \alpha \ln b). \tag{1.47}$$

We have:

$$\exists! \lambda_0 > 0 \quad \textit{such that} \qquad \frac{1}{m}\lambda_0 = \mathcal{F}(\lambda_0, \alpha \ln b). \tag{1.48}$$

In fact, $\lambda \mapsto \mathcal{F}(\lambda, \alpha \ln b)$ is decreasing on $\mathbb{R}^+$and verify

$$\mathcal{F}(0, \alpha \ln b) = b^\alpha \quad \textit{and} \lim_{\lambda \to \infty} \mathcal{F}(\lambda, \alpha \ln b) = 0,$$

so the intermediate value theorem shows that (1.47) *has a unique positive root denoted λ_0. Then one can note that the asymptotic behavior of the solution* (1.45) *is given by the first term of the sum, corresponding to λ_0. We can check that it is exactly what we found previously as the asymptotic behavior (see* (1.23) *and Theorem 1.4.2): by multiple partial integrations we show that definitions* (1.48) *and* (1.22) *of λ_0 are equivalent. In the same way, since* (1.23) *and $U(0) = \lambda_0$, we can establish the following equality giving the asymptotic behavior:*

$$\mathcal{N}(x,t) = C\, e^{\lambda_0 t} N(x) = \frac{a}{m b^\alpha \ln b} \frac{1}{c(\lambda_0)} e^{\lambda_0 t} \frac{1}{x} \left(1 - \frac{\ln x}{\ln b}\right)^{\lambda_0/a - 1}, \tag{1.49}$$

where C is given at (1.36) *and $c(\lambda_0)$ at* (1.46). *Formula* (1.45) *gives us a exponential convergence rate equal to $e^{-(\lambda_0 - \lambda_1)t}$.*

Remark 4 *From* (1.47) *we can deduce the behavior of the $\mu_k = \lambda_k/a$. Let us order them:*

$$\mu_0 > 0 > \mu_1 > \mu_2 > \dots > \mu_k > \dots.$$

We can show that it exists $n_0 \in \mathbb{N}$ such that $\max_{k \neq 0} \mu_k \leq -2n_0 + 1$ and for all $n \geq n_0$ there are exactly two roots μ_n et μ_{n+1} of (1.47) *in $I_n =]-2n, -2n+1[$. Moreover, as n tends to infinity, these roots get closer and closer from the nearest negative integer: $\mu_n + 2n - 1 \to 0$ and $\mu_{n+1} + 2n \to 0$. Finally we have the general behavior of the sequence μ_k:*

$$\mu_k + 2(n_0 - 1) + k \longrightarrow 0. \tag{1.50}$$

This will be confirmed numerically in the following part.

1.5 Numerics

In this section, we discuss numerical difficulties linked to the numerical approximation of the solution of (1.1)-(1.4) and we describe suitable remedies. Equation (1.1) is a conservation law and the first idea to approximate the solution is to use a Godunov scheme. But a naive implementation of this method brings no result. Indeed, without a very fine discretization, which is not affordable for the clinical data ($b = 7.3 \times 10^{10}$), the solution computed that way rapidly blows up.

Furthermore, we keep in mind that we request a sharp description of the large time behavior. Hence, a more adapted strategy should be introduced, that incorporates a subtle treatment of the boundary condition.

In what follows, as detailed in the introduction we choose for w the Gompertzian law (1.8) and we fix

$$\beta(x) = m x^\alpha.$$

1.5.1 Numerical problems linked to the equation and the boundary condition

At first, we note that the variable x varies from 1 to b which is approximately equal to 10^{11} ($b = 7.3 \times 10^{10}$) that makes the use of very fine grids impossible. A change of variable is necessary and we set:

$$y = \ln \frac{b}{x}.$$

But this leads us to consider the following problem

$$\partial_t u - e^y \partial_y(\tilde{w}\, u) = 0,$$

$$u(y, 0) = 0,$$

$$a \ln b\, u(\ln b, t) = \int_0^{\ln b} m b^{\alpha+1} e^{-(\alpha+1)y} u(y, t)\, \mathrm{d}y + \beta(x_p(t)),$$

where $\tilde{w}$ is given by

$$\tilde{w}(y) = aby e^{-y},$$

which is not a conservation law anymore. This remark leads us up to consider the change of function:

$$v(y, t) = b e^{-y} u(y, t),$$

and the final conservation law completed with the following initial and boundary conditions:

$$\partial_t v - \partial_y(ay\, v) = 0, \tag{1.51a}$$

$$v(y, 0) = 0, \tag{1.51b}$$

$$a \ln b\, v(\ln b, t) = \int_0^{\ln b} m b^{\alpha} e^{-\alpha y} v(y, t)\, \mathrm{d}y + \beta(x_p(t)). \tag{1.51c}$$

Secondly, note that in the biological variables the boundary condition is huge compared to the average solution. As we see below, the grid we use in these variable is refined near the boundary (see Figure 1.1): it is adapted to capture the impact of the boundary condition on the solution.

1.5.2 Numerical scheme

We now aim at solving the system (1.51). Since the scales of the solution and the boundary condition are very different for small times, a low order scheme does not approach well the solution. This is why we choose to use the WENO-5 scheme.

It is a finite volume scheme of fifth-order which is based on interpolations of discrete data using polynomials. It is well known that the wider the stencil, the higher the order of accuracy of the interpolation, but this is true only if the interpolated function is smooth inside the stencil. Contrary to traditional finite volume methods, the WENO (Weighted Essentially Non-Oscillatory) scheme does not use fixed stencil interpolations; it uses a convex combination of all candidate stencils instead of the one fixed in traditional schemes. The weights depend non linearly on the smoothness of the interpolated function on each stencil. This scheme is more able than a traditional one to deal with discontinuities of the interpolated function u; it achieves automatically high order accuracy and a non-oscillatory property near discontinuities. See [Rus09] and [Shu98] for more details.

We consider a regular grid in the rescaled variables:

$$0 = y_{\frac{1}{2}} < y_{\frac{3}{2}} < \cdots < y_{J-\frac{1}{2}} < y_{J+\frac{1}{2}} = \ln b.$$

The grid in the biological variables is then very fine near the incoming boundary ($x = 1$), as illustrated in Figure 1.1.

We define the cells and cell centers by

$$I_j = \left[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}} \right],$$

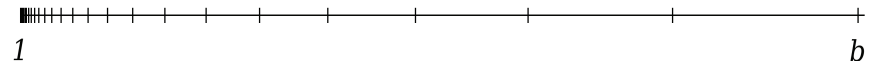


Figure 1.1: Grid displayed in the biological variables (100 points).

$$y_j = \frac{1}{2}\left(y_{j-\frac{1}{2}} + y_{j+\frac{1}{2}}\right),$$
$$\Delta y_i = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}} \equiv h = \frac{\ln b}{J},$$
$$j = 1, 2, \ldots, J.$$

Let

$$\overline{v}_j(t) = \frac{1}{\Delta y}\int_{I_j} v(y,t)\,\mathrm{d}y.$$

Integrating (1.51) on the cell I_j, we get:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{v}_j(t) - \frac{1}{\Delta y}\left[a y_{j+\frac{1}{2}} u(y_{j+\frac{1}{2}}, t) - a y_{j-\frac{1}{2}} u(y_{j-\frac{1}{2}}, t)\right] = 0.$$

Replacing the flux $a y_{j+\frac{1}{2}} u(y_{j+\frac{1}{2}}, t)$ with a monotone numerical flux $\hat{f}(v^-_{j+\frac{1}{2}}, v^+_{j+\frac{1}{2}})$ we get the semi-discretized scheme:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{v}_j(t) - \frac{1}{\Delta y}\left[\hat{f}(v^-_{j+\frac{1}{2}}, v^+_{j+\frac{1}{2}}) - \hat{f}(v^-_{j-\frac{1}{2}}, v^+_{j-\frac{1}{2}})\right] = 0, \tag{1.52}$$

where $v^-_{j+\frac{1}{2}}$ and $v^+_{j+\frac{1}{2}}$ are approximations of the function $v(x,t)$ at the cell boundaries.

WENO reconstruction

On each candidate stencil we reconstruct a 2nd degree polynomial $p_i(x)$:

$$p_2(j) = \frac{1}{3}\overline{v}_{j-2} - \frac{7}{6}\overline{v}_{j-1} + \frac{11}{6}\overline{v}_j,$$
$$p_1(j) = -\frac{1}{6}\overline{v}_{j-1} + \frac{5}{6}\overline{v}_j + \frac{1}{3}\overline{v}_j, \tag{1.53}$$
$$p_0(j) = \frac{1}{3}\overline{v}_j + \frac{5}{6}\overline{v}_{j+1} - \frac{1}{6}\overline{v}_{j+2}.$$

A WENO reconstruction will take a convex combination of the $p_i(j)$, $i \in \{0,1,2\}$, defined in (1.53) as a new approximation of the cell boundary value $v(x_{j+\frac{1}{2}})$:

$$v^-_{j+\frac{1}{2}} = \sum_{r=0}^{2} \omega_r p_r(j) \qquad v^+_{j-\frac{1}{2}} = \sum_{r=0}^{2} \tilde{\omega}_r p_r(j-1).$$

The key of the success of WENO lies in the choice of the weights ω_r. If the function $v(x)$ is smooth in all the candidate stencils, one should take $\omega_r = d_r$ with:

$$d_0 = \frac{3}{10}, \qquad d_1 = \frac{3}{5}, \qquad d_2 = \frac{1}{10}.$$

When the interpolated function has a discontinuity in one or more of the stencils, the corresponding weights have to be almost always 0. To measure the smoothness of the solution, we use smoothness indicators, denoted by β_i. The smaller this indicator β_i, the smoother the function in the target stencil. These β_i are written out explicitly as quadratic forms of the cell averages of v in the stencil:

$$\beta_2(j) = \frac{13}{12}(\overline{v}_{j-2} - 2\overline{v}_{j-1} + \overline{v}_j)^2 + \frac{1}{4}(3\overline{v}_{j-2} - 4\overline{v}_{j-1} + \overline{v}_j)^2,$$

$$\beta_1(j) = \frac{13}{12}(\overline{v}_{j-1} - 2\overline{v}_j + \overline{v}_{j+1})^2 + \frac{1}{4}(\overline{v}_{j-1} - \overline{v}_{j+1})^2,$$

$$\beta_0(j) = \frac{13}{12}(\overline{v}_j - 2\overline{v}_{j+1} + \overline{v}_{j+2})^2 + \frac{1}{4}(\overline{v}_j - 4\overline{v}_{j+1} + 3\overline{v}_{j+2})^2,$$

We then set the nonlinear weights:

$$\omega_r = \frac{\alpha_r}{\alpha_0 + \alpha_1 + \alpha_2},\ r = 0, 1, 2,$$

with

$$\alpha_r = \frac{d_r}{(\epsilon + \beta_r)^2},$$

where ϵ is introduced to prevent the denominator from vanishing. To compute the $\tilde{\omega}_r$ we just note that by symmetry $\tilde{d}_r = d_{2-r}$.

One must note that this scheme needs two additional cells on each side of the integration domain. In our method, we use ghost cells with the boundary values : an approximation of the integral (1.51c) at the boundary $y = \ln b$ and 0 at $y = 0$.

The numerical flux

Concerning the numerical flux, we choose the following Lax-Friedrichs flux:

$$\hat{f}(v^-_{j+\frac{1}{2}}, v^+_{j+\frac{1}{2}}) = \frac{1}{2}\left[-ay_{j+\frac{1}{2}}(v^-_{j+\frac{1}{2}} + v^+_{j+\frac{1}{2}}) - \gamma(v^+_{j+\frac{1}{2}} - v^-_{j+\frac{1}{2}})\right],$$

where $\gamma = \max|-ay| = a\ln b$ is a constant.

To compute this flux, in addition to the value given by the WENO procedure, we need the value of $v^-_{\frac{1}{2}}$ and $v^+_{J+\frac{1}{2}}$. We set $v^-_{\frac{1}{2}} = 0$ and we use the boundary condition to compute $v^+_{J+\frac{1}{2}}$. To this end, we need a robust integration method.

Integration method

In order to compute the integral contained in the boundary condition, we use a Newton-Cotes quadrature formula. Since our WENO reconstruction is of order 5, we need an integration method of order 5 at least. This is why we use Milne's method, the Newton-Cotes quadrature method of order 5, which can be seen also as the Richardson extrapolation adapted to Simpson's method for 3 and 5 points. In order to approach the following integral:

$$I = \int_p^q f(x)\,\mathrm{d}x,$$

we set $h = \frac{q-p}{4}$, $x_i = p + ih$ and $y_i = f(x_i)$. Then

$$J = \frac{2h}{45}\left(7y_0 + 32y_1 + 12y_2 + 32y_3 + 7y_4\right),$$

is an approximation of I of 5th-order.

Note that the Newton-Cotes method needs the value of the function at the nodes of the grid, but our method is a finite volume one. So, starting form the average value of the function on each cell we have to compute the value of the function at the nodes before computing the integral and this has to be of 5th-order. To this end, we can use

again a WENO reconstruction. For the integration we set:

$$
\begin{aligned}
y_{\frac{1}{2}} &= v^+_{\frac{1}{2}},\\
y_{j+\frac{1}{2}} &= \frac{1}{2}\left(v^+_{j+\frac{1}{2}} + v^-_{j+\frac{1}{2}}\right), \quad \text{for } j = 1,\dots,J-1,\\
y_{J+\frac{1}{2}} &= v^-_{J+\frac{1}{2}},
\end{aligned}
$$

where the $v^+_{j+\frac{1}{2}}$, $j = 0,\dots,J-1$, and $v^+_{j+\frac{1}{2}}$, $j = 1,\dots,J$ are given by the WENO procedure, and we perform the integration on this vector y.

Time discretization

We use an explicit high order Runge-Kutta time discretization. The semi-discrete scheme (1.52) written as

$$u_t = L(u),$$

is discretized in time by the Runge-Kutta method of the third order:

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

Remark 5 *As we have seen in Section 2,* (1.51) *is equivalent to the following problem:*

$$
\begin{cases}
\partial_t v - \partial_y(a y\, v) = 0,\\
v(y,0) = \delta(y = \ln b),\\
a \ln b\; v(\ln b, t) = \int_0^{\ln b} m b^\alpha e^{-\alpha y} v(y,t)\,\mathrm{d}y.
\end{cases}
$$

Nevertheless, we choose to implement the first one (1.51)*, because the WENO scheme is able to deal with only one discontinuity per stencil and a Dirac approximation consists of two discontinuities on two successive cells.*

Convergence analysis

Stability We write the stability for the Lax-Friedrichs scheme applied to (1.51). The scheme is described by the following recursion:

$$u_j^{n+1} = \frac{1}{2}\left(1 + \frac{a\Delta t}{\Delta y} y_{j+1}\right) u^n_{j+1} + \frac{1}{2}\left(1 - \frac{a\Delta t}{\Delta y} y_{j-1}\right) u^n_{j-1},$$

$j = 1,\dots,J$, $n \in \mathbb{N}$, where u_j^n stands for the numerical approximation of $v(t^n = n\,\Delta t, y_j = j\,\Delta y)$. Under the CFL condition :

$$\max_{0\le y\le \ln(b)} \left[\frac{\Delta t}{\Delta y} a y\right] \le 1,$$

we get:

$$\begin{aligned}\sum_{j=1}^{J} u_j^{n+1} &\leq \frac{1}{2}\left(1 + a y_{J+1}\frac{\Delta t}{\Delta y}\right) u_{J+1}^n + \sum_{j=1}^{J} u_j^n, \\ &\leq \left(\sup_{0\leq y\leq \ln b} m b^\alpha e^{-\alpha y} + 1\right)\sum_{j=1}^{J} u_j n + m(x_p(t^n))^\alpha, \\ &\leq (m b^\alpha + 1)\sum_{j=1}^{J} u_j^n + m(x_p(t^n))^\alpha,\end{aligned} \tag{1.54}$$

and using Gronwall's lemma we obtain

$$\sum_{j=1}^{J} u_j^n \leq e^{m b^\alpha n}\sum_{j=1}^{J} u_j^0 + \sum_{k=0}^{n-1} m(x_p(t^k))^\alpha e^{m b^\alpha (n-k-1)}.$$

Note that in the homogeneous case, (1.54) gives the L^1 stability.

Consistency Again, we write consistency for the Lax-Friedrichs scheme, and using Taylor developpement we get

$$\left[\frac{v(t^{n+1}, y_j) - \frac{1}{2}(v(t^n, y_{j-1}) + v(t^n, y_{j+1}))}{\Delta t} - \frac{a y_{j+1} v(t^n, y_{j+1}) - a y_{j-1} v(t^n, y_{j-1})}{2\,\Delta y}\right]$$

$$- \left[\partial_t v(t^n, y_j) - \partial_y (a y v)(t^n, y_j)\right]$$

$$= a^2 \frac{\Delta t}{2}\left(\partial_y(y \partial_y (y v)) - \frac{(\Delta y)^2}{a^2 (\Delta t)^2}\partial^2_{y^2} v\right) + O((\Delta y)^2 + (\Delta t)^2).$$

We conclude that Lax-Friedrichs scheme is of 1st-order.

Since we combine a 5th-order WENO reconstruction and a 3rd-order time discretization, we predict that our scheme is convergent of 3rd order. Numerically, we use it to compute the solution of $\partial_t u + \partial_x u = 0$ at $T = 1$, for respectively a gaussian as initial condition $u_0(x) = e^{-4x^2}$ on $[-2; 3]$, and $u_0(x) = \cos(\pi x)$ on $[-1; 1]$ with periodic boundary conditions. The numerical orders we find are respectively 5.14 and 3.92 and our previous argument validated.

1.5.3 Numerical results

In order to validate our scheme, we check that it preserves the asymptotic profile, and then we show that we can derive numerically the Malthusian rate λ_0 from the equation. But first, for experimentation, we need to compute precisely the asymptotic profile, which means essentially to compute the first eigenvalue λ_0 and the constant $c(\lambda_0)$ from (1.47) and (1.46).

For the numerical tests, all the biological constants are taken from [IKS00]:

$$a = 0.00286, \quad \alpha = 0.663, \quad b = 7.3 \times 10^{10}, \quad m = 5.3 \times 10^{-8}.$$

Determination of λ_0 and the constant C defined at (1.36)

In order to evaluate $c(\lambda_0)$, in (1.49), and the λ_k's, we truncate the sums (1.46) and (1.47) at the numerical zero ($\epsilon = 10^{-16}$) and we use the secant method to compute the λ_k's. By analytic considerations, we show (see (1.48)) that there is only one positive root λ_0 of (1.47) which is equal to

$$\lambda_0 = 5.8057 \times 10^{-3}\ days^{-1}. \tag{1.55}$$

λ_0 can be as well determined by the integral formula (1.22). Since the formula defining $c(\lambda_0)$ (see(1.48)) is an alternating sum, one can note that we control the rest of the sum of $c(\lambda_0)$'s formula (1.46) by the first neglected term.

Remark 6 (In reference to Remark 4, page 45.) *Using again the secant method, we can compute the following eigenvalues. More precisely, we want to confirm the asymptotic behavior of the μ_k's suggested in Remark 4. This is consistent with the equivalent given in* (1.50)*. Numerically, we find $n_0 = 25$ and*

$$\begin{aligned}
\mu_1 &= -49.239278160564048,\\
\mu_2 &= -49.8923146858060633,\\
\mu_3 &= -51.0295140700383243,\\
\mu_4 &= -51.990168667673018,\\
\mu_5 &= -53.0030304315224186,\\
\mu_6 &= -53.9990649790580122,\\
\mu_7 &= -55.000281526423727,\\
\mu_8 &= -55.9999165905830267,\\
\mu_9 &= -57.000024264683141,\\
\mu_{10} &= -57.999993061662245.
\end{aligned}$$

Conservation of the asymptotic profile

We impose on our algorithm the asymptotic profile,

$$V(y) = \frac{1}{c(\lambda_0)} \frac{a}{mb^\alpha \ln b} \left(\frac{y}{\ln b}\right)^{\frac{\lambda_0}{a}-1}, \tag{1.56}$$

multiplied by $e^{-\lambda_0 t_{start}}$ as initial condition. For a value of t_{start} large enough, we expect that the solution computed by the algorithm is close to this profile (1.56) growing exponentially with rate λ_0 in time. In fact, it appears that with time the integral in (1.51c) becomes very large compared to the source term which remains bounded, and if we neglect the source term $\beta(x_p(t))$ in (1.51c), with this initial condition (1.56), the exact solution is

$$e^{-\lambda_0 t} b e^{-y} N\left(be^{-y}\right),$$

since $N(\cdot)$ is the eigenvector associated to λ_0, see (1.19a)-(1.19b). At time t_{end}, we compare the computed solution with

$$e^{-\lambda_0 t_{end}} b e^{-y} N\left(be^{-y}\right),$$

(See Figure 1.2). For the test, we choose $t_{start} = 10000\ days$, $t_{end} = 11000\ days$.

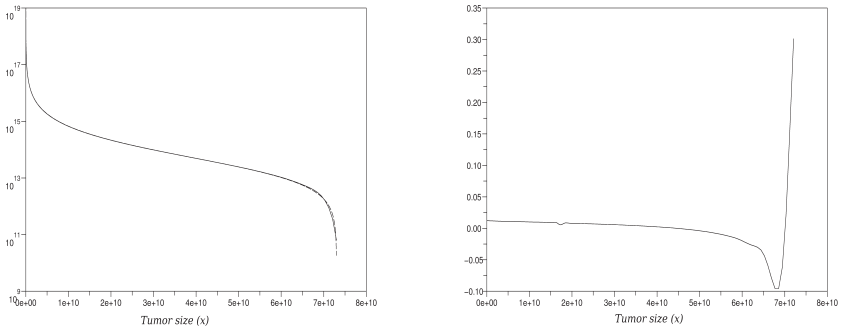


Figure 1.2: On the l.h.s, comparison between the theoretical asymptotic (dashed line) and the computed solution (log scale) and the corresponding relative error on the r.h.s. ; $\alpha = 0.663$, $a = 0.00286$, $b = 7.3 \times 10^{10}$, $m = 5.3 \times 10^{-8}$, $t_{start} = 10000\ days$, $t_{end} = 11000\ days$, $\Delta y = \frac{\ln b}{2000}$.

On Figure 1.2 we plot on the right hand side the relative error. Note that the average relative error is $e_{err} = 2\%$ except for the larger tumors. This difference between the theoretical solution and the computed solution for large tumors can be explained by both the fact the speed w vanishes at $x = b$ and the small number of grid nodes near $x = b$. However, this is exactly what we expect: the asymptotic profile is conserved and it grows at the expected Malthusian rate. Furthermore the asymptotic is conserved whether it takes into account the source term or not; it shows that this conservation property is stable under small perturbations of the boundary condition. We can also observe stability under small perturbations of the initial condition (1.56). All these remarks show that the scheme deals well with the non-local term in the boundary condition.

The Malthusian rate

We come back to the original problem with the initial condition (1.4) and the boundary condition (1.2). Since we showed the asymptotic behavior (1.35), we have

$$I(t) = \int_1^b \rho(x,t)\,\mathrm{d}x \sim_{t\to+\infty} \tilde{C}e^{\lambda_0 t}.$$

We deduce from that remark that the slope of $\ln(I)$ tends to a constant. This can be checked numerically (See Figure 1.3.): if we perform a linear regression on $\ln(I(t))$ for $4500 \leq t \leq 15000$ *days*, we find a slope equal to $s = 5.8039 \times 10^{-3} days^{-1}$ which has to be compared with (1.55). Let us underline that the Malthusian rate is not given by the scheme. It is a characteristic value of the problem that our scheme is able to catch. Again, this shows that our scheme computes a correct approximation of the solution, even for times larger than 10000 days.

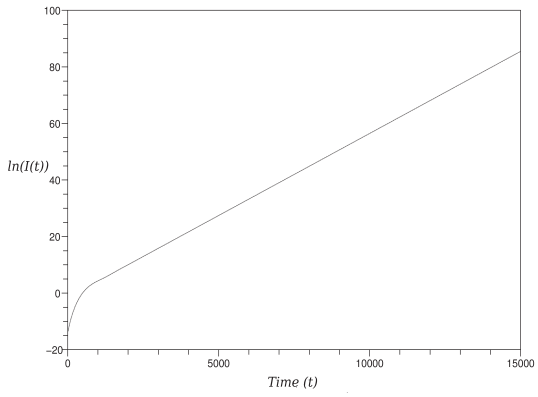
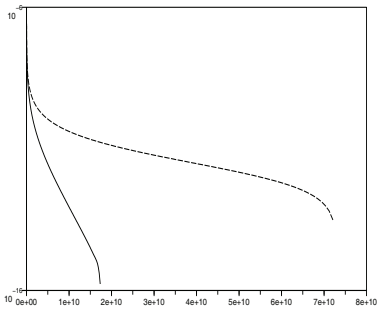


Figure 1.3: $\ln(I(t))$ for $0 \leq t \leq 15000$ *days*, $\alpha = 0.663$, $a = 0.00286$, $b = 7.3 \times 10^{10}$, $m = 5.3 \times 10^{-8}$.

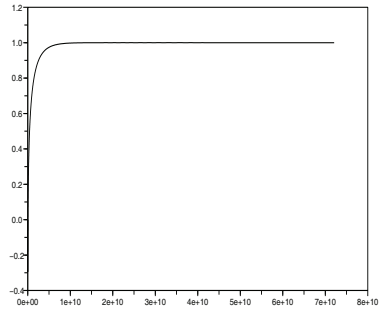
Numerical asymptotic behavior

We plot the numerical solution we obtained with our scheme for several times and we compare it with the theoretical asymptotic behavior given at (1.49). (See Figure 1.4.) Note that the behavior of the computed solution coincides with the asymptotic profile given by the theory for $T = 5000$ *days*. Even for $T = 2000$ *days*, the asymptotic solution is almost reached for small tumors. For the same reasons as in paragraph 1.5.3, the relative error increases again near the boundary $x = b$.

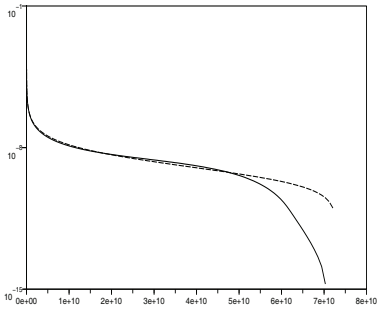
If we perform now the same simulation, just changing the parameter α to 0.4 instead of $\alpha = 0.663$, we reach the U-shape profile predicted in [IKS00]. (See Figure 1.5.) Again, the difference between the theoretical asymptotics and the computed solution for large tumor size can be explained by both the fact that we will always have $\rho(b,t) = 0$, because this tumor size is never reached (see (1.9)) and the small number of grid nodes near $x = b$ (See Figure 1.1).



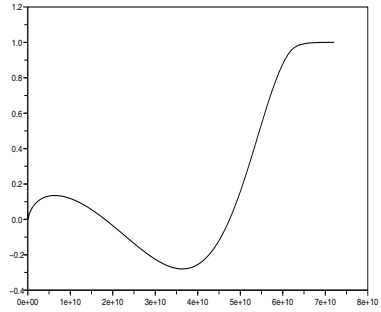
(a) $T = 1000\ days$



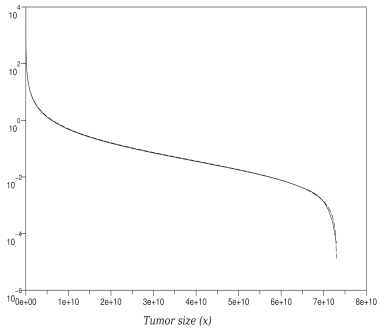
(b) $T = 1000\ days$



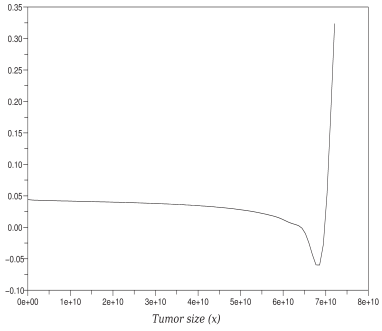
(c) $T = 2000\ days$



(d) $T = 2000\ days$



(e) $T = 5000\ days$



(f) $T = 5000\ days$

Figure 1.4: On the l.h.s, comparison between the theoretical asymptotic (dashed line) and the computed solution (log scale) and the corresponding relative error on the r.h.s. ;
$\alpha = 0.663,\ a = 0.00286,\ b = 7.3 \times 10^{10},$
$m = 5.3 \times 10^{-8},\ \Delta y = \frac{\ln b}{2000}.$

Back to the generational point of view

Intuitively, one can think that only the behaviors of the first descendants of the primary tumor may be sufficient to describe the whole dynamics. Actually this is not right. In order to illustrate it, we compute the 6 first terms of (1.17) and compare them to the theoretical asymptotics. If the first generations are relevant for small times, they are not anylonger after a while. (See Figure 1.6.) A given generation seems to be relevant only on a quite small time interval. For example, the 6-th generation has a relatively small impact on the total behavior of the system for

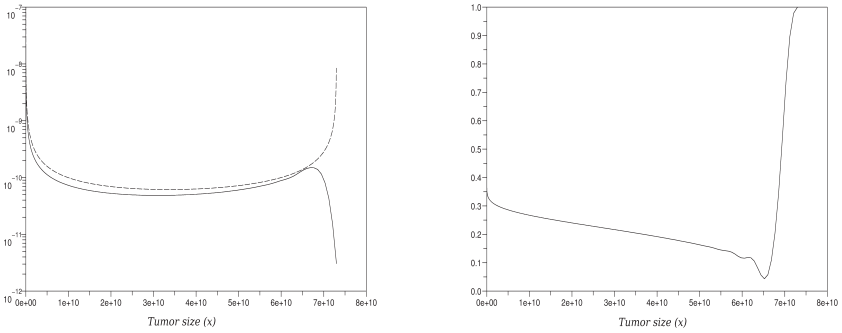


Figure 1.5: On the l.h.s, comparison between the theoretical asymptotic (dashed line) and the computed solution (log scale) and the corresponding relative error on the r.h.s. ; $T = 5000\ days$; $\alpha = 0.4,\ a = 0.00286,\ b = 7.3 \times 10^{10},\ m = 5.3 \times 10^{-8},\ \Delta y = \frac{\ln b}{2000}$.

$t \leq 3500\ days$ and $t \geq 9000\ days$. Each generation has the same growth speed, however to have a relevant impact on the whole population a given generation has to wait for the previous generation to contain enough and large enough tumors to be produced. Next, since the following generation contains more tumors than the previous one (a given tumor could generate more than one malignant single cell), its weight in the whole population will become larger than the previous generation and then will generate more daughter-tumors and so on. This mechanism explains why a given generation has an impact only on a given time interval. However, if we look to the average survival time of a patient, we can consider that for applications, the first generations are relevant.

1.6 Conclusion and perspectives

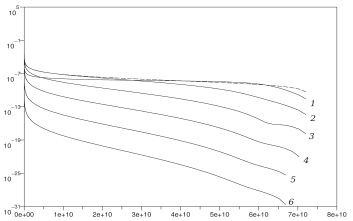
Even if this model does not describe all the complexity of the biological phenomena, it brings to light the existence of a Malthusian parameter which characterizes the exponential growth of metastatic tumors and the long time asymptotics.

A straightforward finite–difference scheme cannot bring relevant results : such a scheme is not able to deal with the large boundary condition at $x = 1$ and the solution we obtain is too large. This is why we propose here a rescaled and a high order scheme. Compared to a characterics scheme, a strategy adopted in [ALM99], [BBHV09], it does not need the expression of the characteristics which are explicit in this simplified model, but will not be anylonger in a more elaborate version.

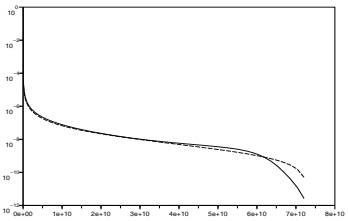
Numerically, we showed that for $T \geq 2000\ days \simeq 5.5\ years$, the theoretical asymptotic profile is a good approximation of the solution. However, from both a practical and a theoretical point of view, the question of the convergence speed to the asymptotic profile is still an open question.

Furthermore, this mathematical model has to be developed. For example, one has to note that no angiogenesis phenomena have been considered, but it is now well known that the tumor does not grow in the same way as before, during and after angiogenic phase of the tumor. These phenomena change the nature of the growth. See [BP00] for a description of the tumor growth phases.

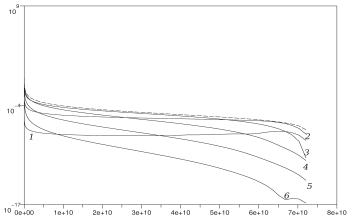
In the future, such models will be helpful for a better understanding of cancer mechanisms. They could be used as well to exhibit small tumors that medical apparatus can not detect and help designing curative strategies.



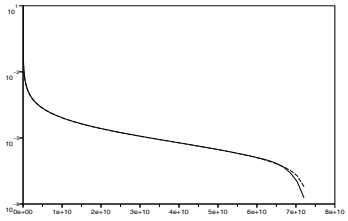
(a) $T = 2300$ *days*.



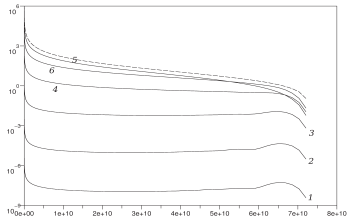
(b) $T = 2300$ *days*.



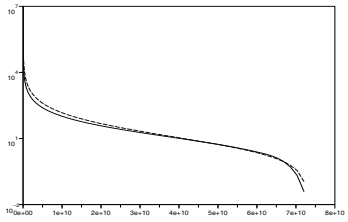
(c) $T = 3500$ *days*.



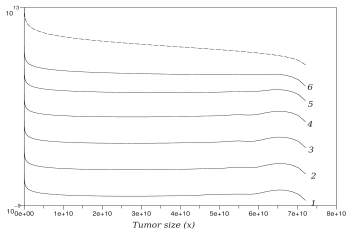
(d) $T = 3500$ *days*.



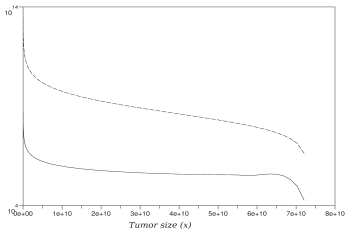
(e) $T = 6000$ *days*.



(f) $T = 6000$ *days*.



(g) $T = 9000$ *days*.



(h) $T = 9000$ *days*.

Figure 1.6: Comparison between the first 6 generations and the theoretical asymptotics (dashed line) for several times on the l.h.s and between the sum of the first 6 generations and the theoretical asymptotics (dashed line) at the same times on the r.h.s. $\alpha = 0.663,\ a = 0.00286,\ b = 7.3 \times 10^{10},$
$m = 5.3 \times 10^{-8}, \Delta y = \frac{\ln b}{2000}$

1.7 Appendix : Discrete model

From a biological point of view, a discrete model can be more understandable than a continuous one. The objective of this part is to understand the problem (1.15) as the limit of a discrete one. First we describe the discrete model and introduce a small parameter, then we give some estimates that are useful to pass to the limit.

1.7.1 The discrete model

We consider a model where the tumors are described by the number of cells they contain. Let c_i be the concentration of tumors containing i cells; c_1 then denotes the concentration of individual tumoral cells. Obviously, we can also think of the index i as characterizing the size of the tumor. The dynamics are governed by the following assumptions:

- the tumors grow with a rate w_i that depends on the number of cells they contain. We choose for w_i a Gompertzian rate ($w_i = a.i \ln\left(\frac{I}{i}\right)$, with I stands for the maximal tumor size and a is a constant). Note that $w_I = 0$.
- the tumors emit individual cells with a rate B_i depending also on the number of cells they contain: the larger this number, the larger the emitting rate. In what follows, we choose $B_i = m.i^\alpha$, where m is a constant and $0 < \alpha \leq 1$.

Accordingly, we are led to the following system of differential equations. For $i \geq 2$, we count the gain of tumors coming from the previous size and the loss due to the growth of the ith-tumors.

$$\frac{\mathrm{d}}{\mathrm{d}t} c_i = w_{i-1} c_{i-1} - w_i c_i, \qquad \text{for } i \geq 2. \tag{1.57}$$

The evolution for the individual cells is slightly different since it counts the loss due to their own growth and the gain due to the production by all tumors; we get

$$\frac{\mathrm{d}}{\mathrm{d}t} c_1 = -w_1 c_1 + \sum_{i=1}^{I} B_i c_i. \tag{1.58}$$

In what follows, we aim at linking this discrete model to the continuous one (1.15) completed with the initial condition:

$$f|_{t=0} = f_0 \tag{1.59}$$

is a given function.

To rescale System (1.57)-(1.58) in terms of a small parameter, let us rewrite it in a dimensionless form. We summarize here reference quantities which will be used in the sequel:

- T: the characteristic time ;
- C_1: the characteristic value for the number of one-cell tumors ;
- C: the characteristic value for the larger tumors, $i \geq 2$;
- B: the characteristic value for the emitting rates B_i;
- W: the characteristic value for the growth rates w_i.

The corresponding dimensionless quantities are defined as:

$$\bar{t} = \frac{t}{T}, \quad \bar{c}_1(\bar{t}) = \frac{c_1(\bar{t}T)}{C_1}, \quad \bar{c}_i(\bar{t}) = \frac{c_i(\bar{t}T)}{C}, \quad \overline{w}_i = \frac{w_i}{W}, \quad \overline{B}_i = \frac{B_i}{B}$$

When one rewrites (1.57)-(1.58) in terms of the new variables, the following dimensionless parameters appear:

$$\gamma = WT, \qquad \text{and} \qquad \kappa = \frac{BCT}{C_1}. \tag{1.60}$$

Omitting the overlines, the system takes the following dimensionless form:

$$\frac{\mathrm{d}}{\mathrm{d}t}c_i = \gamma\,(w_{i-1}c_{i-1} - w_i c_i), \tag{1.61a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}c_1 = -\gamma\, w_1 c_1 + \kappa \sum_{i=1}^{I} B_i c_i. \tag{1.61b}$$

Let us consider small tumors whose size is of $1+h$ order, where h is small. The dimensionless parameters defined in (1.60) and I, the maximal size tumor, appear as coefficients in this problem, and need to be scaled appropriately with respect to h. Let us explain the heuristics of this choice. If the right hand side of the first line of (1.61) is to yield a derivative, then γ should be scaled as the reciprocal of h. To get the equivalent to the integral in the second line, I should be scaled as $(b-1) = Ih$ and $\kappa = 1$.

More precisely, in our study, we consider that initially there are mainly small tumors of size of order $1+h$ and we note that then very large tumors appear only for very long times. This leads us to consider for T of order 1 only tumors which sizes are very small compared to I, the maximal size. Then these tumors "see" I as infinite and by rescaling I relatively to the size of tumors, we have $I = (b-1)/h$. Finally, since we have $\sup_i w_i = aI/e$ the previous considerations show us that W is of order $1/h$. These heuristic remarks lead us to the following choice of the parameters in terms of h:

$$\gamma = \frac{1}{h}, \qquad \kappa = 1, \qquad I = \frac{b-1}{h}.$$

With this choice of the parameters, the rescaled version of (1.61) is

$$\frac{\mathrm{d}}{\mathrm{d}t}c_i^h = \frac{w_{i-1}^h c_{i-1}^h - w_i^h c_i^h}{h}, \qquad \text{for } i \geq 2 \tag{1.62a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}c_1 = -\frac{w_1^h c_1^h}{h} + \sum_{i=1}^{I} B_i^h c_i^h. \tag{1.62b}$$

where

$$h = \frac{b-1}{I},$$

$$w_i^h = a(1+ih)\ln\left(\frac{b}{1+ih}\right)$$

and

$$B_i^h = m(1+ih)^\alpha, \qquad 0 < \alpha \leq 1$$

In the sequel, we denote by f^h the function defined by:

$$f^h(t,x) = \sum_{i=1}^{I} c_i^h(t)\mathbb{1}_{ih+1\leq x<(i+1)h+1},$$

where the c_i^h's are solution of (1.62).

In what follows, we aim to establish the convergence of the rescaled discrete problem to the continuous one (1.15) with (1.59) as $h \to 0$. This work is inspired by [CGPV02]. After establishing *a priori* estimates, we will be able to pass to the limit in (1.62)

Theorem 1.7.1 *Assume that there exist three constants $M_{0,\alpha}$, $M_{0,1}$, $M_{0,2}$ for which $\forall 0 < h < 1$,*

$$\sum_{i=1}^{I} m(1+ih)^\alpha c_i^h(0) \;\leq\; M_{0,\alpha},$$

$$\sum_{i=1}^{I} h c_i^h(0) \;\leq M_{0,1},$$

$$h^2 \sum_{i=0}^{I} i c_i^h(0) \ \leq M_{0,2},$$

Then there exists a subsequence h_n such that:

$$f^{h_n} \rightharpoonup f, \quad \text{in } C^0([0,T]; \mathcal{M}^1([1,b]) - weak - *),$$
$$f_0^{h_n} \rightharpoonup \mu_0, \quad \text{in } \mathcal{M}^1([1,b].$$

where f_0^h is defined by

$$f_0^h(t,x) = \sum_{i=1}^{I} c_i^h(0) \mathbb{1}_{ih+1 \leq x < (i+1)h+1},$$

and where f is the solution of (1.15) *with the initial condition $f\big|_{t=0} = \mu_0$.*

Remark 7 *In particular, if*

$$c_i^h(0) = \frac{1}{h} \int_{1+ih}^{1+(i+1)h} f_0(y)\,\mathrm{d}y$$

for a given function $f_0 \in L^1([1,b])$ then $f_0^h \longrightarrow f_0$ in $L^1([1,b])$ and

$$f^h \rightharpoonup f, \quad \text{in } C^0([0,T]; L^1([1,b]) - weak - *).$$

since we showed that the solution of (1.15) *with a given initial condition is unique.*

1.7.2 Estimates

In this first section we will prove some estimates we will use further on. Note that as a consequence of (1.57)-(1.58), we obtain the following relation for the evolution of the total number of tumors

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{I} c_i^h = \sum_{i=1}^{I} B_i^h c_i^h.$$

More generally, for any sequence φ_i, we get the following weak formulation:

$$\sum_{i=1}^{I} \varphi_i c_i^h(t) = \sum_{i=1}^{I} \varphi_i c_i^h(0) + \int_0^t \sum_{i=1}^{I} \frac{w_i^h c_i^h}{h} (\varphi_{i+1} - \varphi_i)\,\mathrm{d}s + \int_0^t \varphi_1 \sum_{i=1}^{I} B_i^h c_i^h\,\mathrm{d}s. \tag{1.63}$$

Let us at first choose $\varphi_i = (1+ih)^\alpha$. Since $0 < \alpha \leq 1$, we have:

$$0 \leq \varphi_{i+1} - \varphi_i \leq h\alpha(1+ih)^{\alpha-1}.$$

Then, now using (1.63):

$$\begin{aligned}
&\sum_{i=1}^{I} m(1+ih)^\alpha c_i^h(t) \\
&\leq m \sum_{i=1}^{I} (1+ih)^\alpha c_i^h(0) + \int_0^t m \sum_{i=1}^{I} a(1+ih) \ln\left(\frac{b}{1+ih}\right) c_i^h(s) \alpha (1+ih)^{\alpha-1}\,\mathrm{d}s \\
&\quad + m \int_0^t (1+h)^\alpha \sum_{i=1}^{I} m(1+ih)^\alpha c_i^h(s)\,\mathrm{d}s, \\
&\leq m \sum_{i=1}^{I} (1+ih)^\alpha c_i^h(0) + \int_0^t (a\alpha \ln b + 2m) \sum_{i=1}^{I} m(1+ih)^\alpha c_i^h(s)\,\mathrm{d}s,
\end{aligned}$$

(recall that $h \leq 1$), i.e.

$$\sum_{i=1}^{I} B_i^h c_i^h(t) \leq mM_{0,\alpha} + \int_0^t (a\alpha \ln b + 2m) \sum_{i=1}^{I} B_i^h c_i^h(s)\,\mathrm{d}s$$

Using Gronwall's lemma, we get

$$\sup_{0\leq t\leq T} \sum_{i=1}^{I} B_i^h c_i^h(t) \;\leq\; M_B, \tag{1.64}$$

where M_B does not depend on h.

We can now estimate the moments:

- **0th-order moment:** We have:

$$\begin{aligned}
\int_1^b f^h(t,x)\;\mathrm{d}x &= h\sum_{i=1}^{I} c_i^h(t) \\
&= \int_0^t h\sum_{i=1}^{I} B_i^h c_i^h(s)\,\mathrm{d}s + \sum_{i=1}^{I} hc_i(0) \\
&\leq TM_B + M_{0,1} \;\leq\; M^0,
\end{aligned}$$

Finally we get:

$$\sup_{0\leq t\leq T} \int_1^b f^h(t,x)\,\mathrm{d}x \;\leq\; M^0,$$

where M^0 does not depend on h.

- **1st-order moment:** We have:

$$\begin{aligned}
\int_1^b x f^h(t,x)\,\mathrm{d}x &= \sum_{i=1}^{I} \int_{ih+1}^{(i+1)h+1} x c_i^h(t)\,\mathrm{d}x \\
&= h^2 \sum_{i=1}^{I} i c_i^h(t) \;+\; (\frac{h^2}{2} + h) \sum_{i=1}^{I} c_i(t) \\
&= S_1 + S_2.
\end{aligned}$$

S_2 is bounded by $2M^0$ and we have:

$$
\begin{aligned}
S_1 &= h^2 \int_0^t \sum_{i=1}^{I} \frac{w_i^h c_i^h(s)}{h}\,\mathrm{d}s \;+\; \int_0^t h^2 \sum_{i=1}^{I} B_i^h c_i^h(s)\,\mathrm{d}s \;+\; h^2 \sum_{i=0}^{I} i c_i^h(0) \\
&\leq TM_B + M_{0,2} + h^2 \int_0^t \sum_{i=1}^{I} \frac{1}{h} a(ih+1) \ln\left(\frac{b}{ih+1}\right) c_i^h(s)\,\mathrm{d}s \\
&\leq TM_B + M_{0,2} + a \ln b \int_0^t \sum_{i=1}^{I} h^2\, i c_i^h(s)\,\mathrm{d}s + a \ln b\, h \int_0^t \sum_{i=1}^{I} c_i^h(s)\,\mathrm{d}s \\
&\leq TM_B + M_{0,2} + a \ln b \int_0^t h^2 \sum_{i=1}^{I} i c_i^h(s)\,\mathrm{d}s + Ta \ln bM^0 \\
&\leq C + a \ln b \int_0^t h^2 \sum_{i=1}^{I} i c_i^h(s)\,\mathrm{d}s.
\end{aligned}
$$

Using again Gronwall's lemma we get:

$$
\sup_{0\leq t\leq T} \int_1^b x\, f^h(t,x)\,\mathrm{d}x \;\leq\; M^1,
$$

where M^1 does not depend on h.

1.7.3 Proof of Theorem 1.7.1

Let φ be a smooth function ($\varphi \in C^1$) supported in $[1,b]$. We set

$$
\varphi^h(x) = \varphi(x_i) \quad \text{for } x_i = ih+1 \leq x < (i+1)h+1.
$$

Note that φ^h converges uniformly to φ on $[1,b]$.
We fix $\varphi_i = h\,\varphi(ih+1)$ in (1.63). Therefore, we obtain:

$$
\begin{aligned}
\sum_{i=1}^{I} h\varphi(ih+1) c_i^h(t) &= \sum_{i=0}^{I} h\varphi(ih+1) c_i^h(0) \\
&\quad + \int_0^t \sum_{i=1}^{I} \frac{w_i c_i^h(s)}{h} \left(h\varphi((i+1)h+1) - h\varphi(ih+1)\right)\mathrm{d}s \\
&\quad + \int_0^t h\varphi(h+1) \sum_{i=1}^{I} B_i^h c_i^h(s)\,\mathrm{d}s,
\end{aligned}
$$

which is equivalent to :

$$
\begin{aligned}
&\int_1^b f^h(t,x)\varphi^h(x)\,\mathrm{d}x - \int_1^b f^h(0,x)\varphi^h(x)\,\mathrm{d}x \\
&= \int_0^t \int_1^b w^h(x) f^h(\sigma,x) \Delta\varphi^h(x)\,\mathrm{d}x\,\mathrm{d}\sigma \\
&\quad + \int_0^t \varphi(h+1) \int_1^b B^h(x) f^h(\sigma,x)\;\mathrm{d}x\,\mathrm{d}\sigma,
\end{aligned}
\tag{1.65}
$$

where we have set

$$\Delta\varphi^h(x) = \frac{\varphi^h(x+h) - \varphi^h(x)}{h}.$$

It is straightforward to check that as $h \to 0$, the quantity $\Delta\varphi^h$ converges uniformly to φ' on $[1, b]$.

Since $w^h(x)$ is bounded by $(a \ln b)x$ for any x in $[1, b]$ and we have the upperbound (1.64), (1.65) leads to

$$\left| \int_1^b f^h(t_1, x)\varphi^h(x)dx - \int_1^b f^h(t_2, x)\varphi^h(x)dx \right| \leq (Cst.\|\varphi'\|_\infty M^1 + M_B\|\varphi\|_\infty)|t_1 - t_2|$$

which proves that $\left(t \mapsto \int_1^b f^h(t, x)\varphi^h(x)dx\right)_{h\geq 0}$ is equicontinuous on $[0, T]$. It is also equibounded:

$$\left| \int_1^b f^h(t, x)\varphi^h(x)dx \right| \leq \|\varphi\|_\infty M^0.$$

Therefore, by the Ascoli theorem it belongs to a compact set of $C^0([0, T])$.

Moreover, since φ^h approaches φ uniformly on $[1, b]$, one deduces that the sequence $\left(t \mapsto \int_1^b f^h(t, x)\varphi(x)dx\right)_{h\geq 0}$ lies in a compact set of $C^0([0, T])$.
Let φ be in $C_b^0([1, b])$, then by separability , there exists a sequence φ_k of functions in $C^1([1, b])$ such that $\varphi_k \to \varphi$. By the diagonal Cantor process, we can find subsequences, h_n and k_n, and $f \in C^0([0, T], \mathcal{M}^1([1, b]) - \text{weak} - *)$ such that

$$\int_1^b \varphi_{k_n} f^{h_n} \, \mathrm{d}x \to \int_1^b \varphi(x) f(t, \, \mathrm{d}x).$$

But

$$\left| \int_1^b (\varphi - \varphi_{k_n}) f^{h_n}(t, x) \, \mathrm{d}x \right| \leq \|\varphi - \varphi_{k_n}\| M^0,$$

Hence, we can find a subsequence, labeled h_n, and a function f that belongs to $C^0([0, T], \mathcal{M}^1([1, b]) - \text{weak} - *)$ such that the following convergence

$$\int_1^b \varphi(x) f^{h_n}(t, x) \, \mathrm{d}x \to \int_1^b \varphi(x) f(t, \, \mathrm{d}x)$$

as h_n tends to 0, holds uniformly on $[0, T]$, for any $\varphi \in C_b^0([1, b])$ and any $T > 0$.

Now, we are able to pass to the limit in (1.65). Since w and B are regular, w^{h_n} and B^{h_n} converge uniformly to w and B respectively on $[1, b]$. Then, we get for any $\varphi \in C_b^0([1, b])$:

$$\int_1^b f(t, x)\varphi(x) \, \mathrm{d}x \quad - \quad \int_1^b f(0, x)\varphi(x) \, \mathrm{d}x = \int_0^t \int_1^b w(x) f(\sigma, x)\varphi'(x) \, \mathrm{d}x \, \mathrm{d}\sigma$$

$$+ \int_0^t \varphi(1) \int_1^b B(x) f(\sigma, x) \; \mathrm{d}x \, \mathrm{d}\sigma,$$

which correspond to the weak formulation of (1.15) with (1.59). In particular, at $t = 0$, the limit $f(t = 0, x)$ is given by the limit of $f^{h_n}(0, x)$ built from the c_i^{0,h_n}. This completes the proof of the theorem.

1.8 Appendix : Explicit computation of the solution using Laplace's transform

- **Domain of definition:** Since x is bounded, $x \in [0, b]$, we are allowed to suppose the function β is compactly supported.

- **Computation of the characteristics :**

  The characteristics are solution of :

$$\begin{cases} \dfrac{dx_c}{dt} &= g(x) = ax_c(\ln b - \ln x_c), \\ x_c(0) &= x_0. \end{cases}$$

  This system can be solved easily, and we obtain :

$$x_c(t) = b^{1-e^{-at}} b^{\ln(\frac{x_0}{b})e^{-at}} = x_p(t)x_0^{e^{-at}}. \tag{1.66}$$

- ***A priori* upper–bound of the solution:**
  The characteristic linked to our problem are given by (1.66). Along these, we have :

$$\begin{aligned} \frac{d}{dt}\left[\rho(t, x_c(t))\right] &= \frac{\partial \rho}{\partial t}(t, x_c(t)) + \dot{x_c}(t)\frac{\partial \rho}{\partial x}(t, x_c(t)), \\ &= -\frac{\partial (g\rho)}{\partial x}(t, x_c(t)) + g(x_c)\frac{\partial \rho}{\partial x}(t, x_c(t)), \\ &= -\rho(t, x_c(t))\dot{g}(t, x_c(t)); \end{aligned}$$

  thus,

$$\rho(t, x_c(t)) \leq \exp\left(-\int_0^t \dot{g}(x_c(t)\right).$$

  But, we have

$$-a \leq \dot{g}(x_c(t)) \leq a\left(\ln\left(\frac{b}{x_0}\right) - 1\right),$$

  then

$$\boxed{\rho(t, x_c(t)) \leq e^{at}} \tag{1.67}$$

- **Laplace's transform of the equation :**

  The Laplace's transform of ρ reads as follows :

$$\tilde{\rho}(x, s) = \int_0^\infty \rho(x, t)e^{-ts}dt.$$

  Applying this transform to the equation, we get :

$$\int_0^\infty \frac{\partial \rho(x, t)}{\partial t}e^{-ts}dt + \int_0^\infty \frac{\partial}{\partial x}\left(g(x)\rho(x, t)\right)e^{-ts}dt = 0,$$

  then , by intergration by parts :

$$\underbrace{\left[\rho(x, t)e^{-ts}\right]_{t=0}^\infty}_{\substack{\longrightarrow 0 \\ \text{fot } s \text{ large enough}(s > a)}} + s\int_0^\infty \rho(x, t)e^{-ts}dt + \frac{\partial}{\partial x}\left(g(x)\tilde{\rho}(x, s)\right) = 0.$$

We are reduced to solve :

$$\boxed{s\tilde{\rho}(x,s) + \frac{\partial}{\partial x}\left(g(x)\tilde{\rho}(x,s)\right) = 0.} \tag{1.68}$$

- **Resolution of (1.68):**

Equation (1.68) can be solved as follows :

$$\tilde{\rho}(x,s) = \frac{g(1)\tilde{\rho}(1,s)}{g(x)} e^{-s\int_1^x \frac{1}{g(u)}du}.$$

But, according to (1.2)

$$g(1)\rho(1,t) = \int_1^\infty \beta(x)\rho(x,t)dx + \beta(x_p(t)),$$

thus

$$g(1)\tilde{\rho}(1,t) = \int_1^\infty \beta(x)\tilde{\rho}(x,t)dx + \underbrace{\int_0^\infty \beta(x_p(t))e^{-st}dt}_{=B_p(s)}.$$

Let Φ,

$$\Phi(s) = g(1)\tilde{\rho}(1,s) = \int_1^\infty \beta(x)\tilde{\rho}(x,s)dx + B_p(s), \tag{1.69}$$

and then

$$\tilde{\rho}(x,s) = \Phi(s)G(x,s), \tag{1.70}$$

with

$$G(x,s) = \frac{1}{g(x)} e^{-s\int_1^x \frac{1}{g(u)}du}.$$

Pluging (1.70) into (1.69), we get

$$\Phi(s) = \int_1^\infty \beta(x)\Phi(s)G(x,s)dx + B_p(s),$$

which means

$$\Phi(s)\left(1 - \int_1^\infty \beta(x)G(x,s)dx\right) = B_p(s),$$

and then

$$\Phi(s) = \frac{B_p(s)}{1 - \int_1^\infty \beta(x)G(x,s)dx}.$$

Plug this again into (1.70), and get :

$$\boxed{\tilde{\rho} = \frac{B_p(s)G(x,s)}{1 - \int_1^\infty \beta(x)G(x,s)dx}}$$

- **Computation of $G(x,s)$:**
  We compute

$$\begin{aligned}\int_1^x \frac{du}{g(u)} &= \int_1^x \frac{du}{ax\ln\left(\frac{b}{x}\right)},\\ &= -\frac{1}{a}\left[\ln\left(1-\frac{\ln u}{\ln b}\right)\right]_1^x,\\ &= -\frac{1}{a}\ln\left(1-\frac{\ln x}{\ln b}\right),\end{aligned}$$

  thus

$$\boxed{G(x,s)=\frac{1}{ax\ln\left(\frac{b}{x}\right)}\left(1-\frac{\ln x}{\ln b}\right)^{s/a}.}$$

- **Computation of $B_p(s)$:**

$$B_p(s)=\int_0^\infty \beta(x_p(t))e^{-st}dt.$$

  Let the change of variables:

$$\begin{aligned}u &= x_p(t),\\ du &= g(x_p)dt,\\ &= g(u)dt.\end{aligned}$$

  Then, we get

$$\begin{aligned}B_p(s) &= \int_1^b \frac{\beta(u)}{g(u)}\left(1-\frac{\ln u}{\ln b}\right)^{s/a}du,\\ &= \int_1^b \beta(u)G(u,s)du,\\ &= \int_1^\infty \beta(u)G(u,s)du,\end{aligned}$$

  since β is compactly supported. Then, we get:

$$\boxed{B_p(s)=\int_1^\infty \beta(x)G(x,s)dx.}$$

  Let's compute:

$$\begin{aligned}B_p(s) &= \int_0^\infty \beta(x_p(t))e^{-st}dt,\\ &= \int_0^\infty me^{\alpha\left(1-e^{-at}\right)\ln b}e^{-st}dt,\\ &= \int_0^\infty m\sum_{k=0}^\infty \frac{\alpha^k\left(1-e^{-at}\right)^k(\ln b)^k e^{-st}}{k!}dt,\\ &= \sum_{k=0}^\infty \frac{m\alpha^k(\ln b)^k}{k!}\underbrace{\int_0^\infty\left(1-e^{-at}\right)^k e^{-st}dt}_{I_k}.\end{aligned}$$

To compute the I_k's, we look for recurrence equation :

$$\begin{aligned} I_k &= \int_0^\infty \left(1-e^{-at}\right)^k e^{-st}dt, \\ &= \underbrace{\left[\frac{e^{-st}}{s}\left(1-e^{-at}\right)^k\right]_0^\infty}_{=0} + \int_0^\infty \frac{ae^{-at}}{s}k\left(1-e^{-at}\right)^{k-1}e^{-ts}dt, \\ &= \frac{a}{s}k\int_0^\infty e^{-at}\left(1-e^{-at}\right)^{k-1}e^{-ts}dt, \\ &= \frac{a}{s}k\left(\int_0^\infty \left(1-e^{-at}\right)^{k-1}e^{-ts}dt + \int_0^\infty \left(e^{-at}-1\right)\left(1-e^{-at}\right)^{k-1}e^{-ts}dt\right), \\ &= \frac{a}{s}k(I_{k-1}-I_k). \end{aligned}$$

We get finally the following formula :

$$I_k\left(\frac{s}{ak}+1\right) = I_{k-1},$$

with:

$$I_0 = \int_0^\infty e^{-st}dt = \frac{1}{s}.$$

At the end,

$$I_k = \frac{k}{k+\frac{s}{a}}I_{k-1} = \frac{k!}{\left(k+\frac{s}{a}\right)\ldots\left(1+\frac{s}{a}\right)}\frac{1}{s},$$

and

$$\begin{aligned} B_p(s) &= \sum_{k=1}^\infty \frac{m}{s}\frac{(\alpha\ln b)^k}{k!}\frac{k!}{\left(k+\frac{s}{a}\right)\ldots\left(1+\frac{s}{a}\right)} + \frac{m}{s}, \\ &= \frac{m}{s}F\left(1,\frac{s}{a}+1;\alpha\ln b\right), \end{aligned}$$

with

$$F(\alpha,\gamma;z) = \sum_{n=0}^\infty \frac{\alpha(\alpha+1)\ldots(\alpha+n-1)}{\gamma(\gamma+1)\ldots(\gamma+n-1)}\frac{z^n}{n!}, \tag{1.71}$$

where we consider that for $n=0$ the corresponding term equals1. Finally :

$$\boxed{B_p(s) = \int_1^\infty \beta(x)G(x,s)dx = \frac{m}{s}F\left(1,\frac{s}{a}+1;\alpha\ln b\right).}$$

- **Rewriting of $\tilde{\rho}$:** To come back to ρ we need to apply the inverse Laplace's transform to the function $\tilde{\rho}$. To this end, we write it as follows :

$$\begin{aligned} \tilde{\rho}(x,s) &= \frac{B_p(s)G(x,s)}{1-B_p(s)}, \\ &= G(x,s)\left\{\frac{1}{1-B_p(s)}-1\right\}, \\ &= \frac{1}{ax\ln\left(\frac{b}{x}\right)}\left(1-\frac{\ln x}{\ln b}\right)^{s/a}\left\{-1+\frac{s/m}{s/m-F\left(1,\frac{s}{a}+1;\alpha\ln b\right)}\right\}, \\ &= \frac{1}{ax\ln b}\frac{1}{1-\frac{\ln x}{\ln b}}\left(1-\frac{\ln x}{\ln b}\right)^{s/a}\left\{-1+\frac{s/m}{s/m-F\left(1,\frac{s}{a}+1;\alpha\ln b\right)}\right\}. \end{aligned}$$

where:

$$\boxed{\tilde{\rho}(x,s) = \frac{1}{ax\ln b}\left(1-\frac{\ln x}{\ln b}\right)^{s/a-1}\left\{-1+\frac{s/m}{s/m - F\left(1,\frac{s}{a}+1;\alpha\ln b\right)}\right\}.}$$

- **Inverse Laplace's transform :**
  Since we suppose that $\tilde{\rho}(x,s)$ is the result of the Laplace's transform , its poles are bounded by an x_0, and

$$\rho(x,t) = \int_{-\infty}^{+\infty} \tilde{\rho}(x+iy)e^{(x+iy)t}dy, \quad \text{pour} \quad x > x_0 \geq a.$$

  Then we get:

$$\rho(x,t) = \sum_{k=1}^{\infty} Res[\tilde{\rho}(x,s), a\lambda_k],$$

  This leads to

$$\boxed{\rho(x,t) = \frac{1}{ax\ln b}\left(1-\frac{\ln x}{\ln b}\right)^{-1}\sum_{k=1}^{\infty} Res[f(s), a\lambda_k],}$$

  with

$$f(s) = e^{\overbrace{\frac{s}{a}\ln\left(1-\frac{\ln x}{\ln b}\right)+st}^{g(s)}}\left\{-1+\frac{s/m}{s/m - F\left(1,\frac{s}{a}+1;\alpha\ln b\right)}\right\},$$

  where the poles λ_k are the roots of

$$\frac{a}{m}\lambda_k = F(1,\lambda_k+1;\alpha\ln b). \tag{1.72}$$

- **Computation of $Res[f(s), a\lambda_k]$:**
  As in the first part, $f(s)$ is holomorphic, and it reduces to compute the residue of the second factor. We have to compute the limit:

$$\lim_{\lambda\to\lambda_k} a(\lambda-\lambda_k)e^{g(a\lambda)}\frac{a\lambda/m}{a\lambda/m - F(1,\lambda+1;\alpha\ln b)}.$$

  In practive, we look for the limit of the inverse :

$$\begin{aligned} & \lim_{\lambda\to\lambda_k}\frac{m}{a^2}\frac{1}{\lambda(\lambda-\lambda_k)}\left(\frac{a\lambda}{m}-F(1,\lambda+1;\alpha\ln b)\right)e^{-g(a\lambda)}, \\ = \quad & e^{-g(a\lambda_k)}\frac{m}{a^2}\lim_{\lambda\to\lambda_k}\frac{1}{\lambda(\lambda-\lambda_k)}\left(\frac{a\lambda}{m}-F(1,\lambda+1;\alpha\ln b)\right). \end{aligned}$$

But

$$\begin{aligned}
F(1,\lambda+1;\alpha\ln b) &= 1+\sum_{n=1}^{\infty}\frac{(\alpha\ln b)^n}{(\lambda+1)\dots(\lambda+n)},\\
&= 1+\sum_{n=1}^{\infty}(\alpha\ln b)^n\sum_{i=1}^{n}\frac{(-1)^{i-1}}{(i-1)!(n-i)!}\,\frac{1}{\lambda+i},\\
&= 1+\sum_{n=1}^{\infty}(\alpha\ln b)^n\sum_{i=0}^{n-1}\frac{(-1)^{i}}{i!(n-i-1)!}\,\frac{1}{\lambda+i+1},\\
&= 1+\sum_{n=0}^{\infty}(\alpha\ln b)^{n+1}\sum_{i=0}^{n}\frac{(-1)^{i}}{i!(n-i)!}\,\frac{1}{\lambda+i+1},\\
&= 1+\sum_{n=0}^{\infty}(\alpha\ln b)\sum_{i=0}^{n}\frac{(-\alpha\ln b)^{i}}{(\lambda+i+1)i!}\,\frac{(\alpha\ln b)^{(n-i)}}{(n-i)!},\\
&= 1+(\alpha\ln b)b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^n}{(\lambda+n+1)n!},\\
&= 1-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda+n+1)n!}.
\end{aligned}$$

From this formulation of F, we deduce :

$$\frac{\partial}{\partial\lambda}F(1,\lambda+1;\alpha\ln b)=b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda+n+1)^2n!},$$

and then

$$\begin{aligned}
&\lim_{\lambda\to\lambda_k}\frac{1}{\lambda(\lambda-\lambda_k)}\left(\frac{a\lambda}{m}-F(1,\lambda+1;\alpha\ln b)\right),\\
=\ &\lim_{\lambda\to\lambda_k}\frac{1}{\lambda}\left(\frac{a}{m}-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda_k+n+1)^2n!}\right),\\
=\ &\frac{1}{\lambda_k}\left(\frac{a}{m}-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda_k+n+1)^2n!}\right),\\
=\ &\frac{1}{\lambda_k^2}\left(\frac{a\lambda_k}{m}-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{n!}\,\frac{\lambda_k}{(\lambda_k+n+1)^2}\right),\\
=\ &\frac{1}{\lambda_k^2}\left(1-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda_k+n+1)n!}-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{n!}\,\frac{\lambda_k}{(\lambda_k+n+1)^2}\right),\\
=\ &\frac{1}{\lambda_k^2}\left(b^{\alpha}b^{-\alpha}-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda_k+n+1)^2(n+1)!}\left((n+1)(\lambda_k+n+1)+\lambda_k(n+1)\right)\right),\\
=\ &\frac{1}{\lambda_k^2}\left(b^{\alpha}+b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(n+1)!}-b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda_k+n+1)^2(n+1)!}\left(2\lambda_k(n+1)+(n+1)^2\right)\right),\\
=\ &\frac{1}{\lambda_k^2}\left(b^{\alpha}+b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^{n+1}}{(\lambda_k+n+1)^2(n-1)!}\left((\lambda_k+n+1)^2-2\lambda_k(n+1)-(n+1)^2\right)\right),\\
=\ &b^{\alpha}\sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^n}{n!(\lambda+n)^2},\\
=\ &b^{\alpha}c(\lambda_k).
\end{aligned}$$

Finally we get :

$$\boxed{\rho(x,t) = \frac{a}{mb^{\alpha}\ln b}\,\frac{1}{x}\sum_{k=1}^{\infty} e^{a\lambda_k t}\left(1-\frac{\ln x}{\ln b}\right)^{\lambda_k-1}\frac{1}{c(\lambda_k)}.} \tag{1.73}$$

with:

$$c(\lambda_k) = \sum_{n=0}^{\infty}\frac{(-\alpha\ln b)^n}{n!(\lambda_k+n)^2}.$$

Acknowledgments

We thank D. Barbolosi who pointed us Ref. [IKS00] and introduced us to the problem. During the preparation of this work we learnt a lot from many fruitful discussions with B. Perthame, J. Clairambault and E. Grenier. We thank also P. Michel for his fruitful remarks. We are also gratefully indebted to A. Ben Abdallah, F. Hubert and F. Verga for friendly exchanges on the different approaches to the problem.

Bibliography

[ALM99] Oscar Angulo and J. C. López-Marcos.
Numerical schemes for size-structured population equations.
Math. Biosci., 157(1-2):169–188, 1999.
Deterministic models with applications in population dynamics and other fields of biology (Sofia, 1997).

[BBHV09] Dominique Barbolosi, Assia Benabdallah, Florence Hubert, and Federico Verga.
Mathematical and numerical analysis for a model of growing metastatic tumors.
Math. Biosci., 218(1):1–14, 2009.

[BP00] Nicola Bellomo and Luidgi Preziosi.
Modelling and mathematical problems related to tumor evolution and its interaction with the immune system.
Math. Comput. Modelling, 32(3-4):413–452, 2000.

[CGPV02] Jean-Francois Collet, Thierry Goudon, Frédéric Poupaud, and Alexis Vasseur.
The Beker-Döring system and its Lifshitz-Slyozov limit.
SIAM J. Appl. Math., 62(5):1488–1500 (electronic), 2002.

[DGL09] Anne Devys, Thierry Goudon, and Pauline Lafitte.
A model describing the growth and the size distribution of multiple metastatic tumors.
DCDS-B, 12(4):731–767, 2009.

[FM86] Chris L. Frenzen and James D. Murray.
A cell kinetics justification for Gompertz' equation.
SIAM J. Appl. Math., 46(4):614–629, 1986.

[GP06] Piotr Gwiazda and Benoît Perthame.
Invariants and exponential rate of convergence to steady state in the renewal equation.
Markov Process. Related Fields, 12(2):413–424, 2006.

[GW89] Mats Gyllenberg and Glenn F. Webb.
Quiescence as an explanation of gompertzian tumor growth.
Growth Development & Aging, 53:25–33, 1989.

[GW90] Mats Gyllenberg and Glenn F. Webb.
A nonlinear structured population model of tumor growth with quiescence.
J. Math. Biol., 28(6):671–694, 1990.

[IKS00] Kazuro Iwata, Kohkichi Kawasaki, and Nanako Shigesada.
A dynamical model for the growth and size distribution of multiple metastatic tumors.
Journal of Theoretical Biology, 203, Issue 2:177–186, March 2000.

[KB03] Frank Kozusko and Željko Bajzer.
Combining Gompertzian growth and cell population dynamics.
Math. Biosci., 185(2):153–167, 2003.

[Mic05] Philippe Michel.
Principe d'entropie généralisée et dynamique de populations structurées.
Thèse de doctorat, Université Paris IX-Dauphine, november 2005.

[MMP04] Philippe Michel, Stéphane Mischler, and Benoît Perthame.
General entropy equations for structured population models and scattering.
C. R. Math. Acad. Sci. Paris, 338(9):697–702, 2004.

[Per07] Benoît Perthame.
Transport equations in biology.
Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2007.

[Rus09] Giovanni Russo.
High order shock capturing schemes for balance laws.
In *Numerical solutions of partial differential equations.*, pages 160–251. Advanced Courses in Mathematics - CRM Barcelona. Basel: Birkhäuser. , 2009.

[Shu98] Chi-Wang Shu.
Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws.
In *Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997)*, volume 1697 of *Lecture Notes in Math.*, pages 325–432. Springer, Berlin, 1998.

Deuxième partie

Modélisation de la respiration

Chapitre 2

Le poumon humain

Dans ce chapitre, nous donnons une brève description de l'appareil respiratoire pulmonaire. Nous nous attachons à donner les outils physiologiques et mécaniques nécessaires pour comprendre la modélisation des phénomènes étudiés dans les Chapitres 3 et 4. Le but est de donner les informations et l'intuition nécessaire à la compréhension des modèles présentés, non de faire une étude exhaustive de la physiologie pulmonaire. Le lecteur pourra s'il le souhaite passer ce chapitre et venir s'y référer si besoin au cours de sa lecture des Chapitres 3 et 4.

2.1 Introduction

Notre poumon possède une structure extraordinaire. Le plus grand organe de notre corps, d'un volume de 7 litres environ, contient un demi–litre de sang, moins d'un demi–litre de tissus, le reste étant de l'air. Sa structure en arbre offre dans plus de 300 millions d'alvéoles une surface d'échange de l'ordre de 100 à 140 m^2. Le mécanisme associé à cet organe est la respiration. Sa fonction est d'apporter de l'oxygène aux tissus et d'évacuer le dioxyde de carbone. Cette tâche peut être détaillée en 4 fonctions principales :

i) *la ventilation pulmonaire* qui concerne les flux d'air entre l'atmosphère et les alvéoles pulmonaires ;

ii) *la diffusion de l'oxygène et du dioxyde de carbone entre les alvéoles et le sang* ;

iii) *le transport de l'oxygène et du dioxyde de carbone* vers ou depuis les tissus cellulaires ;

iv) *la régulation de la ventilation*.

Dans ce qui suit nous nous intéresserons essentiellement aux deux premiers points. En effet, une première partie du travail (Chapitre 3) concerne la simulation des flux d'air dans l'arbre bronchique, tandis que la seconde (Chapitre 4) concerne la modélisation de l'absorption de l'oxygène par le système respiratoire.

Ce chapitre s'organise comme suit. Nous donnerons tout d'abord quelques indications concernant la physiologie de la respiration en commençant par décrire de la mécanique musculaire, puis les voies respiratoires proprement dites en donnant notamment deux différents modèles pour sa géométrie. Nous détaillerons ensuite les données physiologiques que l'on peut obtenir à l'aide d'un examen de spirométrie, examen souvent prescrit chez les patients atteints de pneumopathies. Les résultats de ces examens sont importants car ils constituent les rares données physiologiques nous permettant de valider nos modèles.

Après avoir détaillé l'absorption de l'oxygène à la surface d'échange, nous donnerons quelques notions d'hydrodynamique en précisant les domaines de validité des modèles qui seront utilisés par la suite (modèles de Navier–Stokes et de Poiseuille).

Nous terminerons ce chapitre en présentant le modèle mécanique d'équation différentielle ordinaire (dit 0D) qui sera couplé à un modèle fluide (Chapitre 3) ou à un modèle pour la concentration en oxygène (Chapitre 4).

2.2 Physiologie de la respiration

2.2.1 Mécanique de la ventilation

Les voies respiratoires sont enchassées dans le *parenchyme pulmonaire*, ensemble de tissus contenant les sacs alvéolaires, le système vasculaire pulmonaire et les muscles pulmonaires. Le tout est enfermé dans une boîte formée par la cage thoracique sur les côtés et par le *diaphragme*, vaste et mince nappe musculaire séparant la cage thoracique et les poumons de la cavité abdominale (Figure 2.1).

Le poumon s'élargit et se contracte de deux manières :

i) par le mouvement vertical du diaphragme, rallongeant ou raccourcissant la cavité thoracique ;

ii) par l'ouverture ou la fermeture de la cage thoracique par action des muscles intercostaux augmentant ou diminuant le diamètre de la cage thoracique.

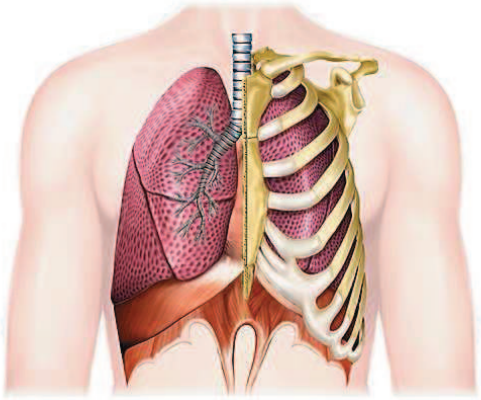


FIG. 2.1: Le poumon dans la cage thoracique. Source : `www.larousse.fr`

Lors d'une respiration normale au repos, la respiration est essentiellement dirigée par le mouvement du diaphragme. À l'inspiration, la contraction du diaphragme tire la partie inférieure du poumon, créant une dépression permettant l'entrée d'air frais dans les poumons. À l'expiration, le diaphragme se relâche, et le système retourne à l'équilibre sous l'action de la force de retour élastique, expulsant ainsi l'air des poumons. À l'effort, la force de retour élastique n'étant pas suffisante pour obtenir une expiration efficace et rapide, une force supplémentaire est exercée à l'expiration par les muscles abdominaux, compressant un peu plus les poumons.

2.2.2 Géométrie des voies respiratoires

Les voies respiratoires sont formées de "tubes" s'organisant en forme d'arbre conduisant à des sacs appelés *alvéoles*, où a lieu l'échange gazeux. La forme arborée conduit naturellement à parler de générations en notant "génération 0" la branche mère constituée de la trachée. On compte ainsi 24 générations, trachée incluse, que nous numérotons de 0 à 23. (Voir Figure 2.2.) On distingue souvent deux régions :

1. les 15 premières générations sont purement convectives : elles n'ont pour fonction que le transport de l'air. On appelle cette première partie la *région proximale*.

2. Après une zone de transition apparaît à partir de la 17e génération la *région distale* : l'air y est bien ralenti et on voit apparaître le long des bronches des *acini* formant des grappes de petites poches, les *alvéoles*.

Étant donné la progression géométrique dans l'arbre, il faut garder à l'esprit que les alvéoles constituent la plus grande partie du volume pulmonaire et la surface d'échange atteint une taille comparable à celle d'un demi terrain de tennis.

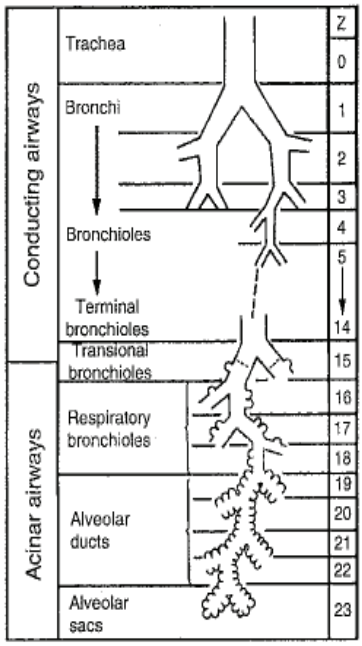


FIG. 2.2: Vue schématique de l'arbre bronchique [WEIBEL]

Modélisation de l'arbre bronchique Le rôle de l'arbre respiratoire est d'apporter l'air frais, riche en oxygène vers la surface d'échange avec le sang, au niveau des alvéoles. Le transport de l'air dans l'arbre bronchique a fait l'objet de nombreux travaux : citons par exemple ceux de MAUROY [Mau04] et SOUALAH [SA07]. Il nous

faut aussi citer ici l'ouvrage de référence de WEIBEL [Wei84]. Dans chacune de ces références on trouve une description détaillée de l'arbre pulmonaire.

Sa structure prend la forme d'un arbre dichotomique, chaque branche se divisant en deux et ainsi de suite, lui donnant ainsi une dimension fractale. Bien que les conduits aient une forme légèrement conique, pour la modélisation nous les supposerons dans toute la suite cylindriques. Du fait des variations dans la population, et de la difficulté à mesurer concrètement les caractéristiques géométriques de l'arbre trachéobronchique, il existe divers modèles apportant des données quantitatives pour la géométrie. Nous en utiliserons deux :

1. **Le modèle homothétique** où chacune des branches filles est la copie homothétique de sa mère, formant un arbre *auto–similaire*. Dans ce cas, la longueur de la trachée est fixée à $L_0 = 12\text{cm}$, son le diamètre à $D_0 = 1.8\text{cm}$ et le rapport d'homothétie à $\rho = 0.85$ [Mau04]. Ainsi, toute la géométrie de l'arbre peut être retrouvée par homothétie. La valeur du rapport d'homothétie peut varier selon les références. En effet, WEIBEL montre que le facteur optimal d'un point de vue hydrodynamique correspond approximativement à la racine cubique de $1/2$ [WG62]. Cependant, une analyse plus récente et approfondie [Mau04, MFWS04] montre que le facteur de réduction est un peu plus grand : 0.85 au lieu de $0.79 \backsimeq \sqrt[3]{1/2}$. Ce léger excès est à comprendre comme une sécurité vis–à–vis de l'optimum. En effet, dans ces références les auteurs montrent qu'une construction optimisée de l'arbre bronchique (avec un facteur égal à 0.79) pourrait être dangereuse car une très légère constriction des bronches ou des bronchioles induirait une augmentation massive de la résistance à la ventilation, qui croît avec la puissance 4 du diamètre, voir Équation (2.7), page 82.
2. **Le modèle de Weibel** [Wei84, WSF05]. Dans les faits, les premières bifurcations ne sont pas symétriques car le poumon gauche doit laisser de la place au coeur. Cela s'uniformise peu à peu et on peut raisonnablement parler d'arbre dichotomique homothétique à partir de la 5e ou 6e génération. Jusqu'à la génération 15, l'arbre peut être considéré comme dichotomique et auto–similaire. Pour les 8 dernières générations la longueur des bronches semble stagner. Nous indiquons dans la Table 5.1, page 135, les valeurs utilisées [Wei84, Mau04] dans le cadre de ce modèle.

2.2.3 Physiologie des voies respiratoires

En termes de physiologie l'ouvrage de GUYTON & HALL [GH06] est une mine d'informations. Pour plus de détails le lecteur peut aussi consulter l'ouvrage de WEST [Wes79].

Mucus et muscle lisse

Toutes les voies respiratoires, depuis le nez jusqu'aux bronchioles terminales sont maintenues humides par un mucus qui recouvre entièrement la paroi interne. Plus que de maintenir les surfaces humides, le mucus piège les petites particules apportées par l'air qui n'atteindront ainsi jamais les alvéoles. Le mucus est lui même régulièrement évacué de la façon suivante : il est remonté continûment par des *cils* puis au niveau du pharynx est dégluti. Pour cette raison nous pourrons supposer par la suite que l'air adhère parfaitement aux parois bronchiques.

Tandis que les bronches du haut (*i.e.* jusqu'à la 5e génération) sont formées essentiellement de cartilage et donc sont relativement rigides, pour les générations suivantes, les parois sont essentiellement formées par le *muscle lisse*. Ce muscle permet aux bronchioles de se déformer en modifiant leur diamètre au cours de la respiration. Ce muscle disparaît pratiquement à partir de la 16e génération où apparaissent les alvéoles formées essentiellement d'épithélium (=tissu) pulmonaire. Dans la suite la déformation des bronches sera modélisée par un paramètre θ distribuant le volume d'air entrant dans les bronches et les alvéoles suivant l'activité du muscle lisse.

Volumes pulmonaires : espace mort et capacités pulmonaires

En respiration normale, l'organisme n'utilise pas toute sa capacité pulmonaire. Afin d'étudier la ventilation pulmonaire, on utilise un spiromètre qui enregistre les volumes d'air entrant et sortant au cours de la respiration (voir page 76 pour plus de détails sur la spirométrie). Ce type de mesure donne le diagramme de la Figure 2.3. On découpe le volume total du poumon en 4 sous–volumes :

1. **le volume courant** (*tidal volume*) : le volume d'air inspiré et expiré lors d'une respiration normale. Chez un jeune homme, ce volume est de l'ordre de 0.5 L au repos ;

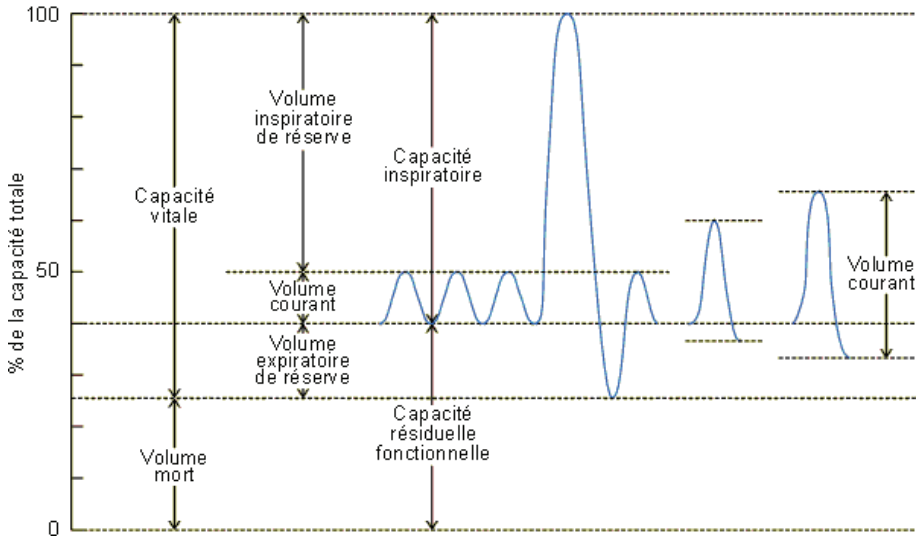


FIG. 2.3: Diagramme présentant les variations de volume pulmonaire durant une respiration normale, inspiration et expiration maximales, Source : `http ://coproweb.free.fr`.

2. **la capacité inspiratoire de réserve** : le volume d'air supplémentaire pouvant être inspiré en plus du volume courant lorsque la personne exécute une inspiration forcée. Il représente environ un volume de 3 L ;
3. **la capacité expiratoire de réserve** : le volume d'air maximal pouvant être expiré lors d'une expiration forcée en plus du volume courant. Ce volume est estimé à 1.1 L ;
4. **le volume mort** : le volume d'air restant dans les poumons après une expiration maximale. Cela correspond environ à $1,2$ L d'air.

Espace mort et renouvellement de l'air alvéolaire. Une partie de l'air inspiré n'atteint jamais les alvéoles où a lieu l'échange gazeux, il remplit simplement les premières bronches où aucun échange gazeux n'est possible tels le nez, la trachée... Cet air occupe l'*espace mort*, c'est–à–dire qu'il ne joue aucun rôle dans l'échange gazeux.

À l'expiration, l'air de l'espace mort est expiré en premier, bien avant que l'air alvéolaire ne soit expiré ; on comprend que l'espace mort est un handicap au renouvellement de l'air alvéolaire. Dans les faits, en respiration normale, à chaque cycle seulement $1/7^e$ de l'air alvéolaire est renouvelé par de l'air frais [GH06].

Un examen : courbe de spirométrie

La spirométrie est la méthode de référence pour dépister les pathologies pulmonaires, l'asthme par exemple. Il existe plusieurs tests de spirométrie, le plus important est la *Capacité Vitale Forcée* (CVF).

Description du test. Le patient exécute une manoeuvre en respirant à travers un embout relié à la machine. Il commence par respirer calmement dans le spiromètre. Lorsque le patient et l'appareil sont prêts, le patient inspire, profondément afin de remplir ses poumons à fond, et expire ensuite le plus fort et le plus vite possible tout l'air de ses poumons dans le spiromètre. L'expiration doit durer entre 3 et 6 secondes. Enfin le patient inspire complètement et forcé afin d'obtenir la courbe inspiratoire. Le spiromètre nous fournit deux courbes : la courbe temps–volume et la courbe débit–volume, qu'on appellera aussi *portrait de phase*. Il est impressionnant de voir la quantité d'informations que peut extraire un spécialiste de ce genre de courbe.

Interprétation des résultats. Nous donnons dans ce paragraphe quelques éléments pour interpréter les courbes débit–volume. Sur une courbe débit–volume, la partie supérieure correspond à l'expiration et la partie inférieure à l'inspiration (voir Figure 2.4). Durant la manoeuvre la courbe se parcourt dans le sens horaire en partant de l'origine. Pour l'expiration, la première partie (à gauche du graphe) correspond au début de l'expiration forcée,

avec le débit expiratoire de pointe (DEP). L'air expiré durant cette période est l'air situé dans les bronches les plus hautes. Plus on se déplace vers la droite sur le graphe, plus l'air expiré correspond à celui des bronches les plus basses. On comprend donc par exemple qu'un défaut dans la courbe situé au niveau du pic correspond à une pathologie dans les bronches hautes tandis que une courbe trop concave correspondrait à une pathologie des petites bronches.

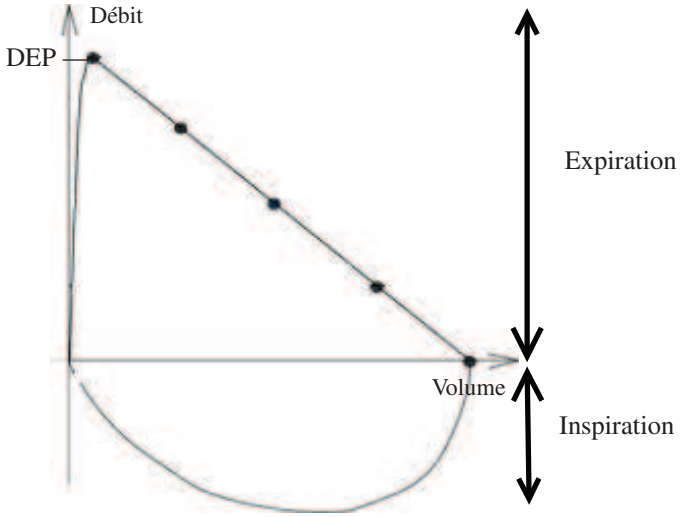


FIG. 2.4: La courbe débit–volume.

Afin de rendre les choses concrètes, nous représentons à la Figure 2.5 quelques croquis de courbes débit–volume dans les cas des pathologies les plus importantes : le *syndrome obstructif* et le *syndrome restrictif*. À la Figure 2.5(a) nous donnons l'allure de la courbe dans le cas d'une spirométrie normale à titre de comparaison.

- **Syndrome obstructif :** La Figure 2.5(b) représente un syndrome obstructif. Chez les patients ayant un tel syndrome, les petites bronches sont partiellement obstruées : c'est ce qui se passe en cas d'asthme[3] ou de broncho-pneumopathie chronique obstructive (BPCO ou bronchite chronique)[4]. Dans ce cas la courbe débit–volume est caractérisée par une concavité. L'air des hautes voies peut être expiré correctement, le débit de pointe est donc normal. Après l'expiration de l'air des voies hautes, l'air expulsé correspond à celui des petites bronches. Si ces bronches sont obstruées, l'air sort moins vite, et le débit est diminué. C'est pour cela qu'on observe une courbe située en dessous de la courbe « normale » et concave.

- **Syndrome restrictif :** La Figure 2.5(c), représente un syndrome restrictif, ce qui correspond à un volume pulmonaire trop faible. Les voies respiratoires étant normales, la courbe débit–volume a une allure normale, simplement « raccourcie » le long de l'axe représentant le volume. Un tel syndrome est observé chez des patients atteints de fibrose[5] Au contraire, chez les patients atteints d'emphysème,[6] on observe une augmentation de la capacité totale, donc un écart selon l'axe des abscisses plus grand.

- **Syndrome mixte :** Souvent des patients présentent les deux syndromes présentés ci–dessus. La Figure 2.5(d), représente une courbe débit–volume montrant les caractéristiques des deux syndromes.

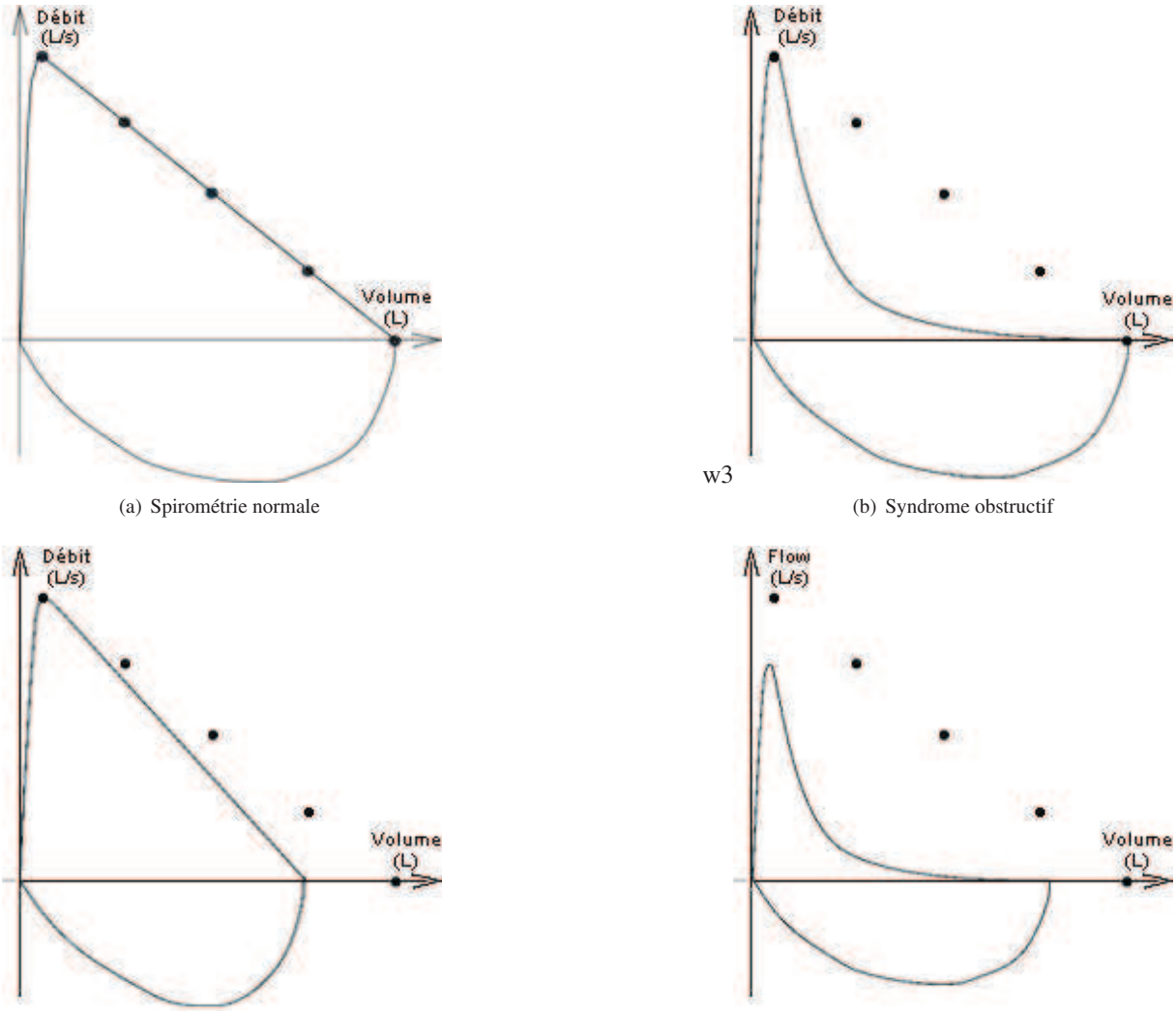
Pour terminer, en guise d'illustration nous présentons à la Figure 2.6 un véritable résultat d'examen pour un patient asthmatique.

[3]Gêne respiratoire à l'expiration, causée par une obstruction des bronches. Les causes peuvent être multiples et il existe différents types d'asthmes

[4]Obstruction d'origine infectieuse virale ou bactérienne

[5]Altération chronique et progressive du tissu pulmonaire (parenchyme) dont les alvéoles sont enserrées et étouffées par des fibres.

[6]Maladie des alvéoles pulmonaires, définie par l'augmentation de volume (dilatation) des alvéoles pulmonaires avec destruction de leur paroi élastique, ce qui entraîne l'impossibilité pour elles de se vider complètement, à l'expiration, de l'air qu'elles contiennent.



(a) Spirométrie normale

(b) Syndrome obstructif

(c) Syndrome restrictif

(d) Syndrome mixte

FIG. 2.5: Interprétation de la courbe débit–volume. Source : `www.spirometrie.info`

2.2.4 Les échanges gazeux

La concentration en oxygène est décroissante le long de l'arbre : effet de « screening »

L'absorption d'oxygène est inhomogène le long de l'arbre bronchique ; en effet on observe un appauvrissement de l'air en oxygène à mesure que l'on descend dans l'arbre bronchique. Afin qu'un acinus soit efficace, l'oxygène doit pouvoir atteindre les bronches terminales, malgré l'appauvrissement en oxygène. Dans la plupart des modèles, notamment tous les modèles mono–compartiment [BT06, Bat09], chaque bronche se termine par une chambre en forme de bulle de telle façon que tous les échanges gazeux aient les mêmes caractéristiques. Cependant, on sait que les poches d'échange gazeux sont plutôt arrangées comme des chambres latérales appelées alvéoles, qui apparaissent le long des bronches. La ventilation des alvéoles le long de l'arbre est donc plutôt en série (voir Figure 2.2), certaines sont proches de l'entrée d'air (*i.e.* de la bouche et du nez) d'autres sont vraiment à la périphérie. L'efficacité des échanges gazeux dépend de l'approvisionnement en air frais, et donc de la situation de l'avéole dans l'arbre bronchique. Notons que du fait de la géométrie dichotomique de l'arbre, la moitié de la surface totale d'échange se situe sur la dernière génération.

Nous notons deux problèmes majeurs qui affectent l'efficacité des échanges gazeux dans le poumon, problèmes

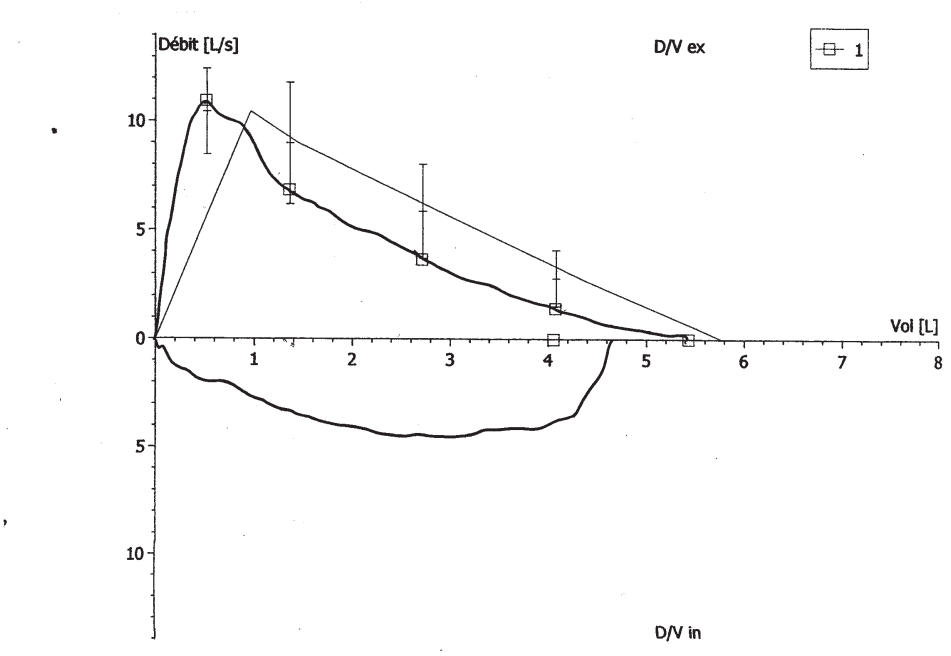


FIG. 2.6: Un résultat d'examen par spirométrie

que nous voulons prendre en compte par la suite dans notre modèle :

a) L'arbre bronchique a la géométrie d'un arbre quasi–dichotomique. Pour un arbre homothétique (voir Page 75), il est facile de voir que la section totale croît exponentiellement au fur et à mesure qu'on s'enfonce dans l'arbre. Les relevés morpho–métriques montrent qu'il en est de même dans le cas d'un arbre réel. En conséquence, la vitesse du fluide décroît et passe au bout d'un certain temps en dessous de la vitesse de diffusion de l'oxygène. Cette transition a lieu approximativement à l'entrée des acini, mais cela dépend fortement du régime respiratoire : le point de transition convection/diffusion se déplace vers le fond de l'arbre lorsqu'on passe d'un régime de respiration de repos à celui de l'effort. Cette transition est mesurée par le *nombre de Péclet* (voir page 82).

b) Tandis que l'oxygène diffuse le long de l'arbre, il est peu à peu absorbé à travers la surface d'échange disposée le long de son parcours ce qui a pour conséquence la diminution de la pression partielle en oxygène. Ce phénomène connu sous le nom de *« screening effect »* (effets d'écrantage) (voir [WSF05, HGC$^+$05] par exemple) est important car il détermine quelle quantité d'oxygène atteint les alvéoles situées au plus profond de l'arbre. Il y a compétition entre deux phénomènes : dans la partie haute de l'acinus, la pression partielle en oxygène est importante mais la surface d'échange est faible. Au contraire, au fond de l'acinus, la surface d'échange est très importante mais du fait des effets de screening la pression en oxygène est faible.

Diffusion de l'oxygène de la chambre alvéolaire vers le sang

L'alvéole est entouré de capillaires sanguins (Figure 2.7). Les alvéoles sont groupés par 5 ou 6 pour former un acinus pulmonaire desservi par une microbronchiole. Elles sont le lieu des échanges gazeux du poumon. Ainsi, c'est dans l'alvéole que l'air donne une partie de son oxygène pour transformer le sang veineux rouge sombre en sang artériel rouge vif. L'air contenu dans l'alvéole s'enrichit alors en gaz carbonique, évacué à l'expiration. Pour passer de la chambre alvéolaire vers le sang, l'oxygène doit traverser une membrane. Le fraction passant dans le sang est liée à plusieurs paramètres notamment :

- la différence de pression partielle en oxygène entre le sang et l'alvéole. La pression partielle en oxygène dans le sang à son arrivée dans le système vasculaire pulmonaire est de 40mmHg [GH06] ;
- le ratio perfusion/ventilation qui mesure le rapport entre la fréquence respiratoire et la fréquence cardiaque, qui varie selon le régime respiratoire. D'après [GH06], ce ratio est estimé à $1/6$ au repos et $1/2$ à l'effort ;

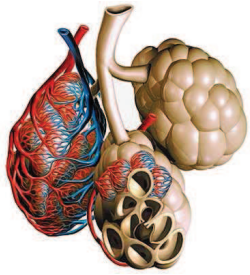


FIG. 2.7: Acini pulmonaires. Source : `www.larousse.fr`

- la perméabilité de l'oxygène à travers la membrane, qui est constante ;
- la surface d'échange efficace, à savoir la surface alimentée en air frais.

2.3 Mécanique de la respiration

2.3.1 Un peu d'hydrodynamique

Les équations de Navier–Stokes

Le modèle hydrodynamique de Navier–Stokes [Bat70] propose la description de l'écoulement d'un fluide visqueux dans un domaine via sa vitesse $\mathbf{u}(t,\mathbf{x})$ et sa pression $p(t,\mathbf{x})$. Ces équations correspondent aux lois physiques de conservation de la masse, de la quantité de mouvement et de l'énergie. Dans le cas de l'air dans les poumons, nous faisons en plus l'hypothèse que le fluide est Newtonien incompressible. Ce qui nous amène au système suivant :

$$\left\{\begin{array}{rcll} \rho\dfrac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \eta\Delta\mathbf{u} + \nabla p &=& \mathbf{F_{ext}} & t\in\mathbb{R}_+,\quad \mathbf{x}\in\Omega \\ \nabla\cdot\mathbf{u} &=& 0 & t\in\mathbb{R}_+,\quad \mathbf{x}\in\Omega, \end{array}\right. \tag{2.1}$$

où ρ est la masse volumique de l'air, supposée dans toute la suite constante, et η le coefficient de viscosité dynamique. $\mathbf{F_{ext}}$ correspond à l'action éventuelle de forces extérieures sur le fluide, par exemple la gravité. La première équation de (2.1) traduit la conservation de la quantité de mouvement, tandis que la seconde impose l'incompressibilité du fluide.

L'hypothèse d'incompressibilité se justifie par un *nombre de Mach* faible. Le nombre de Mach, noté *Ma* exprime le rapport de la vitesse locale d'un fluide sur la vitesse du son dans ce même fluide. Dans le cas du poumon, la vitesse de l'air dans la trachée, lors de la respiration au repos, est de l'ordre de $1\,\mathrm{m}\cdot\mathrm{s}^{-1}$ diminuant le long de l'arbre (de l'ordre de $0.1\,\mathrm{m}\cdot\mathrm{s}^{-1}$ dans la 5e génération), valeur négligeable comparé à la vitesse du son dans l'air ($340\,\mathrm{m}\cdot\mathrm{s}^{-1}$). Généralement, on peut négliger les variations de masse volumique lorsque le nombre de Mach vérifie $Ma < 0.3$; c'est largement le cas ici et cela le reste également dans le cadre du respiration forcée pour laquelle le nombre de Mach est de l'ordre de $0.05 - 0.08$.

De plus, nous verrons au paragraphe suivant que le nombre de Reynolds étant relativement peu élevé, l'utilisation du modèle d'Euler, considérant l'air comme un fluide parfait non visqueux, est exclue.

Nombre de Reynolds

Le terme non linéaire de convection $(\mathbf{u}\cdot\nabla)\mathbf{u}$ dans les équations de Navier–Stokes est sans doute celui qui pose le plus de problèmes tant sur le plan physique, que mathématique ou numérique. Dans certains cas, nous pouvons supposer ce terme négligeable ce qui simplifie grandement le problème. Sur quel critère pouvons–nous justifier d'une telle approximation ? L'idée est de comparer le terme d'inertie $\rho\left[\dfrac{\partial\mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u}\right]$ au terme de forces volumiques de viscosité $\eta\Delta\mathbf{u}$. Soit U (resp. L) une vitesse caractéristique pour le fluide (resp. une longueur caractéristique du

domaine), alors, en utilisant les variables adimensionnées suivantes :

$$\bar{\mathbf{x}} := \frac{\mathbf{x}}{L}, \quad \bar{\mathbf{u}}(\bar{t}, \bar{\mathbf{x}}) := \frac{\mathbf{u}(t, \mathbf{x})}{U}, \quad \bar{t} := \frac{U}{L} t, \quad \bar{p}(\bar{t}, \bar{\mathbf{x}}) := \frac{L}{\eta U},$$

le système de Navier–Stokes (2.1) peut se réécrire de la manière suivante :

$$\left\{ \begin{array}{rcll} \dfrac{\partial \bar{\mathbf{u}}}{\partial \bar{t}} + (\bar{\mathbf{u}} \cdot \nabla)\bar{\mathbf{u}} + \dfrac{\eta}{\rho U L} \left(-\eta \Delta \bar{\mathbf{u}} + \nabla \bar{p}\right) & = & \bar{\mathbf{F}}_{\mathbf{ext}}, & \bar{t} \in \mathbb{R}_+, \quad \bar{\mathbf{x}} \in \bar{\Omega}, \\ \nabla \cdot \bar{\mathbf{u}} & = & 0, & \bar{t} \in \mathbb{R}_+, \quad \bar{\mathbf{x}} \in \bar{\Omega}, \end{array} \right. \tag{2.2}$$

où apparaît le nombre sans dimension

$$\mathrm{Re} := \frac{\rho U L}{\eta}, \tag{2.3}$$

appelé *nombre de Reynolds*. Le nombre de Reynolds caractérise l'écoulement étudié. Il existe un seuil au delà duquel au observe un brusque changement de régime. Ce seuil dépend fortement de la géométrie. Pour un tube par exemple, on peut considérer que pour un nombre de Reynolds inférieur à 2300, les lignes de courant fluide sont toutes parallèles, caractéristique d'un régime *laminaire*. Si on dépasse cette valeur critique, le régime est dit *turbulent* et des tourbillons apparaissent. Lorsque le nombre de Reynolds est petit, inférieur à l'unité, on peut alors négliger les termes d'inertie et on obtient le système de Stokes suivant :

$$\left\{ \begin{array}{rcll} -\eta \Delta \mathbf{u} + \nabla p & = & \mathbf{F}_{\mathbf{ext}}, & t \in \mathbb{R}_+, \quad \mathbf{x} \in \Omega, \\ \nabla \cdot \mathbf{u} & = & 0, & t \in \mathbb{R}_+, \quad \mathbf{x} \in \Omega, \end{array} \right. \tag{2.4}$$

où nous avons omis les barres.

Estimons le nombre de Reynolds dans la trachée à partir de la définition (2.3) ; pour une vitesse caractéristique de $1\,\mathrm{m} \cdot \mathrm{s}^{-1}$ et une longueur caractéristique de l'ordre du centimètre, on obtient un Reynolds autour de 500. Toutefois, si on considère plutôt une vitesse de l'ordre de $1\,\mathrm{m} \cdot \mathrm{s}^{-1}$ et une longueur caractéristique de 2 centimètres (diamètre de la trachée) le nombre de Reynolds est de l'ordre de 2000. Ainsi, localement, dans les parties hautes du moins, le nombre de Reynolds peut être localement plus élevé. La vitesse et la longueur caractéristiques de la géométrie étant toutes les deux décroissantes le long de l'arbre, nous remarquons que le nombre de Reynolds décroît également le long de l'arbre. Ce nombre va nous permettre de diviser la partie proximale du poumon en deux pour construire le modèle qui nous intéresse au Chapitre 3. En effet, alors que le comportement de l'air dans la partie supérieure sera décrit par les équations de Navier–Stokes, dans la seconde, le nombre de Reynolds étant petit devant 1, nous utiliserons alors le modèle de Stokes, caractérisé dans un tube par l'écoulement de Poiseuille.

Remarque 1 *La plupart des auteurs considère que, lors de la respiration, l'écoulement reste laminaire [FMP+05] car le phénomène en présence est oscilatoire ne laissant pas le temps aux instabilités de se développer (le nombre caractéristique est alors le Womersley). Cependant dans [LTMH07] des simulations numériques mettent en évidence que la géométrie complexe du larynx peut induire un écoulement turbulent dans le cas d'un débit constant rentrant dans les voies respiratoires. Notons toutefois que les équations de Navier-Stokes décrivent non seulement les écoulements laminaires mais également les écoulements turbulents (la difficulté est alors de capturer les petites échelles lors des simulations numériques).*

Écoulement de Poiseuille

L'écoulement de Poiseuille décrit le comportement d'un fluide non–inertiel dans un tube. Son domaine d'utilisation est large : cela va de la conduite de chauffage aux vaisseaux sanguins. Comme mentionné précédemment, le flux d'air dans le poumon à partir de la 7e génération peut être condidéré comme non–inertiel ; dans ce cas de fluide lent parcourant des conduits de petits diamètres, il est possible d'obtenir les profils de vitesse et de pression comme solutions explicites du problème de Stokes.

Loi de Poiseuille tridimensionnelle Plaçons nous dans un cylindre de diamètre D de longueur L et d'axe (O, x). Nous supposons qu'il y circule un fluide incompressible non–inertiel vérifiant les équations de Stokes (2.4) avec des conditions de non–glissement sur la paroi, *i.e.* $\mathbf{u} = 0$ sur la paroi. Par symétrie évidente la vitesse du fluide

est invariante par rotation d'axe (O,x). Nous notons donc $\mathbf{u}(r,x)$ la vitesse du fluide situé à l'abscisse x et à une distance r de l'axe du cylindre. Nous notons P_0 la pression supposée uniforme à l'entrée du tube, et P_L la pression à la sortie. La vitesse et la pression sont alors données par :

$$\begin{aligned} \mathbf{u} &= \frac{(P_0 - P_L)D^2}{16\eta L}\left(1 - \frac{4r^2}{D^2}\right)\mathbf{e}_x, \\ p &= -\frac{P_0 - P_L}{L}x + P_0, \end{aligned} \tag{2.5}$$

où $\mathbf{e}_x$ est un vecteur directeur unitaire de l'axe (O,x). On note que la vitesse ne dépend pas de la coordonnée x et la pression, linéaire en x, est constante sur chaque section du tube. Remarquons aussi le profil parabolique de la vitesse, qui prend sa valeur maximale au centre du tube. Par intégration nous trouvons la forme suivante pour le débit, c'est–à–dire le volume de fluide traversant une section transverse par unité de temps :

$$Q = \frac{\pi D^4}{128\eta L}(P_0 - P_L). \tag{2.6}$$

Cette formule, appelée *loi de Poiseuille*, donne le débit comme proportionnel à la différence de pression. Nous notons

$$R = \frac{128\eta L}{\pi D^4}. \tag{2.7}$$

Par analogie avec les lois de l'électricité où l'intensité est proportionnelle à la différence de potentiel (*loi d'Ohm*), nous pouvons calculer la résistance équivalente d'un réseau de tuyaux. Nous remplaçons ainsi par exemple dans le cas du poumon, une partie de l'arbre bronchique par un tube de résistance équivalente.

Loi de Poiseuille bidimensionnelle Au Chapitre 3, nous réalisons des simulations numériques à l'aide du logiciel de calcul éléments finis `Freefem++` en dimension $d = 2$, où le domaine de calcul est une partie du plan. Nous pouvons réécrire la loi de Poiseuille dans ce cas et obtenons :

$$R = \frac{12\eta L}{2D^3}. \tag{2.8}$$

Remarque 2 *Au Chapitre 3 les calculs sont effectués en dimension $d = 2$ pour une géométrie homothétique (voir page 75). Le rapport d'homothétie est alors choisi de façon à ce que les résistances 2D obtenues correspondent aux résistances réelles 3D. Voir [SA07].*

Vitesse du fluide et nombre de Peclet

Sur les 5 premières générations, la vitesse de l'air varie entre 1 et $0.1\mathrm{m}\cdot\mathrm{s}^{-1}$. Puis la vitesse diminue, et les phénomènes de diffusion prennent le pas sur la convection.

La grandeur mesurant le rapport entre la convection et la diffusion est le *nombre de Péclet*. Il est défini comme suit :

$$Pe := \frac{LU}{D_{O_2}}, \tag{2.9}$$

où L est la distance parcourue caractéristique (la longueur de la bronche), U la vitesse de fluide caractéristique (la valeur moyenne de la vitesse du fluide dans la bronche) et D_{O_2} le coefficient de diffusion de l'oxygène dans l'air. Le calcul du nombre de Péclet revient à faire le rapport entre la distance moyenne parcourue par convection et la distance moyenne parcourue par diffusion sur un temps donné. Si $Pe >> 1$ alors les effets convectifs seront prédominants, tandis que si $Pe << 1$ le transport se fera essentiellement par diffusion.

En régime de respiration normale, la limite convection–diffusion se situe au niveau des générations 17–18. À l'effort cette barrière descend, et se situe au niveau des générations 21–22.

Remarque 3 *Les nombres de Péclet et de Reynolds sont proportionnels ; on a l'identité suivante :*

$$Pe = Re \cdot Sc,$$

où Sc désigne le *nombre de Schmidt* qui représente le rapport entre la diffusivité de quantité de mouvement $\nu = \frac{\eta}{\rho}$ (ou viscosité cinématique) et de la diffusivité massique D_{O_2}. Il est utilisé pour caractériser les écoulements de fluides dans lesquels interviennent simultanément viscosité et transfert de matière.

2.4 Modèles couplés

2.4.1 Le modèle mécanique 0D

Le modèle mécanique que nous utilisons dans toute la suite est un modèle possédant un unique degré de liberté x correspondant à l'amplitude du thorax. Ce modèle a déjà été décrit notamment dans [Bat09] et utilisé dans [BGM09, MMSS08]. Dans ce modèle le poumon est assimilé à une boîte, dont une face est mobile (voir Figure 2.8). La face mobile modélise le parenchyme ; elle est reliée à un ressort, représentant l'élasticité des tissus.

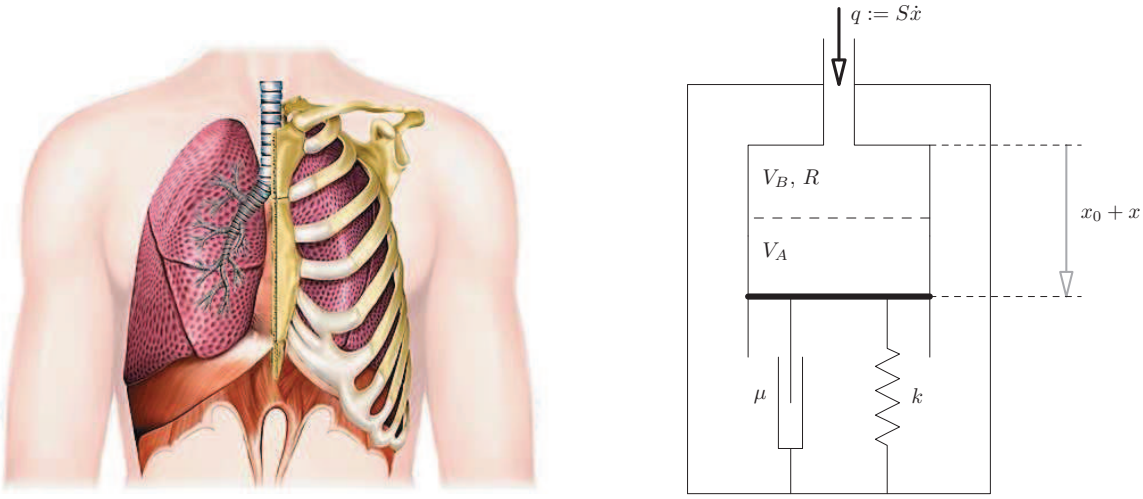


FIG. 2.8: Modèle mécanique : le poumon est modélisé comme une boîte dont une paroi est mobile. $V = S\,(x_0 + x)$: volume total du poumon, R : résistance totale des voies aériennes, k : constante de raideur du ressort, μ : coefficient de friction. La partie mobile a pour masse m et pour surface S. Le volume peut être divisé en deux parties : l'espace bronchique (de volume V_B et de résistance R) et l'espace alvéolaire (de volume V_A).

Paramètres

Nous introduisons ici des paramètres et variables entrant en jeu dans notre modèle :

- x le déplacement du piston par rapport à sa position d'équilibre. Physiologiquement, cette grandeur est à relier au déplacement des côtes et du parenchyme ;
- m la masse du poumon, ou plus précisément la masse déplacée de sorte que $m\dot{x}^2/2$ corresponde à l'énergie cinétique mise en jeu ;
- k_0 la constante de raideur du ressort, qui mesure la raideur relative aux forces élastiques tendant à ramener naturellement le poumon à sa position d'équilibre ;
- S la surface du ressort correspondant à la surface effective du parenchyme sur laquelle la pression due à la force s'applique ;
- μ la résistance des tissus, de sorte que $\mu\dot{x}$ représente le travail dissipé dans les tissus ; la résistance des tissus résulte des frictions entre le poumon et la cage thoracique ;
- R_0 mesure la résistance de l'arbre bronchique de sorte que la différence entre la pression extérieure et la pression intérieure soit égale à R_0 multiplié par le flux, la pression extérieure étant la pression atmosphérique et la pression intérieure, la pression alvéolaire. La résistance des voies aériennes est conséquence de la friction des molécules d'air contre les parois ;
- f_{ext} représente les forces extérieures, conséquences de l'effort musculaire produit pour la déformation de la cage thoracique induisant l'inspiration et l'expiration.

Notons p la pression alvéolaire et normalisons à zéros la pression atmosphérique. Puisque le flux correspond à la dérivée du volume, nous avons l'identité suivante : $\Delta p = -p = R_0 S\dot{x}$. En appliquant le principe fondamental de la dynamique à la partie mobile, nous obtenons l'EDO suivante :

$$m\ddot{x} + (R_0 S^2 + \mu)\dot{x} + k_0 x = f_{ext}.$$

Remarquons ici que la résistance totale est la mise en série de deux résistances : la résistance des bronches R_0 et la résistance du poumon $R_\ell = \mu/S^2$. Au paragraphe suivant, nous prendrons en compte des effets non–linéaires pour la résistance R et la constante de raideur du ressort k ; c'est pour cela que l'on va plutôt considérer l'EDO suivante :

$$m\ddot{x} + (R(x)S^2 + \mu)\dot{x} + k(x)x = f_{ext}, \tag{2.10}$$

où $k(x)$ et $R(x)$ sont précisés ci–après.

On donne à l'Annexe 5.2 et 5.3 page 136, les valeurs des paramètres du fluide et du système masse-ressort utilisées pour les calculs.

Variations de la constante de raideur du ressort dues à la contrainte de volume critique

Dans le cas d'une manoeuvre forcée, lors d'une profonde inspiration ou expiration, le volume de la cage thoracique ne peut pas excéder ou passer en dessous d'une certaine taille critique. Ce modèle a été proposé originellement dans [MMSS08]. Cette limitation peut être prise en compte dans notre modèle en considérant une constante de raideur du ressort linéaire entrainant une non–linéarité dans le modèle. Nous notons $\delta V_{max} = Sx_{max} > 0$ (resp. $\delta V_{min} = Sx_{min} < 0$) la variation de volume maximale (resp. minimale) déterminée expérimentalement. Durant une manoeuvre forcée, nous autorisons le poumon du patient à atteindre ces volumes critiques, et nous devons supposer la constante de raideur de ressort non–linéaire. Fixons :

$$k(x) := k_0 + k^{(e)}(x) + k^{(i)}(x),$$

où $k^{(e)}(x)$ (resp. $k^{(i)}(x)$) est la contribution du ressort à l'expiration (resp. à l'inspiration) c'est–à–dire pour x négatif (resp. positif). Supposons :

$$k^{(e)}(x) := \begin{cases} K_e\, x/x_{min}, & \text{si} \quad x < 0 \\ 0, & \text{si} \quad x \geq 0 \end{cases}, \qquad k^{(i)}(x) := \begin{cases} 0, & \text{si} \quad x \leq 0 \\ K_i\, x/x_{max}, & \text{si} \quad x > 0 \end{cases},$$

où K_e et K_i sont des constantes positives à préciser. Notons que la constante de raideur croît linéairement aussi bien lorsque le volume pulmonaire croît que lorsqu'il décroît, empêchant ainsi le volume d'atteindre des valeurs au delà des valeurs critiques. Remarquons aussi que pour de petites variations de volume, *i.e.* pour x petit, la non–linéarité n'a que peu d'influence sur le système. Les constantes K_e et K_i sont estimées comme suit : situons–nous lors d'une manoeuvre forcée à l'expiration, ce qui signifie qu'une force maximale négative $f_{ext} = f_{min} < 0$ est exercée. Dans ce cas, comme nous pouvons l'observer sur des courbes expérimentales nous avons $\ddot{x} = \dot{x} = 0$ et l'équation (2.10) se réduit à

$$k(x_{min})x_{min} = f_{min},$$

ce qui nous donne immédiatement $K_e = f_{min}/x_{min} - k_0$. De manière analogue, en se plaçant à l'inspiration maximale, nous obtenons $K_i = f_{max}/x_{max} - k_0$. Finalement, le modèle non–linéaire pour la constante de raideur du ressort se réécrit de la façon suivante :

$$k(x) = k_0 + \begin{cases} (f_{min}/x_{min} - k_0)x/x_{min}, & \text{si} \quad x < 0, \\ (f_{max}/x_{max} - k_0)x/x_{max}, & \text{si} \quad x \geq 0. \end{cases} \tag{2.11}$$

Remarque 4 *Dans le cas d'une respiration au repos, où le déplacement x reste relativement faible, le modèle donne $k(x) \sim k_0$.*

Variations de la résistance des voies aériennes dues aux propriétés mécaniques de l'arbre bronchique

On suit encore le modèle de [MMSS08], afin de prendre en compte l'influence du muscle lisse (voir page 75). Essentiellement le muscle lisse empêche les bronches de trop se déformer. Dans notre modèle, nous devons supposer que la variation totale du volume d'air pulmonaire se distribue entre les alvéoles et les bronches suivant un paramètre de distribution θ. La variation totale de volume $\delta V = Sx$ est donc égale à la somme de deux termes : δV_A pour les alvéoles et δV_B pour les bronches avec

$$\delta V_A = (1-\theta)Sx, \qquad \delta V_B = \theta Sx.$$

Cette distribution est simplement fondée sur la conservation du volume, en considérant le volume des tissus (parenchyme) constant. En notant V_A^0 (resp. V_B^0) le volume alvéolaire (resp. bronchique) au repos et V_A (resp. V_B) le volume alvéolaire (resp. bronchique) courant, la relation précédente se réécrit

$$V_A = V_A^0 + (1-\theta)Sx, \qquad V_B = V_B^0 + \theta Sx.$$

En supposant, que les bronches se déforment tout en gardant leur forme originale, supposée cylindrique, et puisque la résistance d'une bronche est proportionnelle à L/D^4 (voir page 81), où L correspond à la longueur de la bronche et D au diamètre de celle-ci, nous obtenons que la résistance est inversement proportionnelle au volume. C'est pourquoi nous choisissons pour la loi de comportement de la résistance des voies aériennes la loi suivante :

$$R(x) = \frac{R_0}{1 + \theta Sx/V_B^0}. \tag{2.12}$$

On donne les valeurs numériques pour le calcul de la constante de raideur et de la résistance dans les Annexes 5.4 et 5.5 page 136.

Comparaison avec les données physiologiques

Afin de valider notre modèle mécanique, nous présentons des simulations dans les deux cas : au repos et à l'exercice, en utilisant pour les forces extérieures les valeurs données en Annexe 5.6, page 137. La seule grandeur pour laquelle la comparaison est réellement possible est sur le calcul du volume inspiré au cours de chaque cycle. La Table 2.1 donne les résultats numériques obtenus par simulation de l'Équation (2.10) et les données que l'on peut trouver dans la littérature [GH06, Wei84].

	modèle mathématique	données physiologiques
repos	0.58 L	0.475 L [Wei84]
effort	2.5 L	2.27 L [Wei84]

TAB. 2.1: Comparaison des volumes inspirés au repos et à l'effort donnés par le modèle mathématique aux volumes physiologiques.

Les résultats sont en bon accord avec les données physiologiques : nous obtenons les bons ordres de grandeur. Nous ne cherchons pas à être beaucoup plus précis, étant donné la variabilité dans la population.

2.4.2 Un modèle multi–échelle pour l'appareil respiratoire

Dans la première partie de ce travail (Chapitre 3), nous nous intéressons à un modèle multi–échelle décrivant le flux d'air dans les poumons. Nous nous attachons dans cette section à décrire le modèle en question. Ce modèle suit une décomposition en trois parties du poumon qui correspondent chacune à un modèle mécanique différent.

Comme nous l'avons vu, la respiration nécessite le transport de l'air, à travers l'appareil respiratoire dont les extrémités visibles sont le nez et la bouche. À l'autre extrémité de l'arbre bronchique on trouve les alvéoles nichés dans un tissu viscoélastique, le tout enfermé dans une boîte formée latéralement par la cage thoracique, et en dessous par le diaphragme. L'air est mis en mouvement via le déplacement du diaphragme et du réseau tissulaire qui y est rattaché, le *parenchyme*.

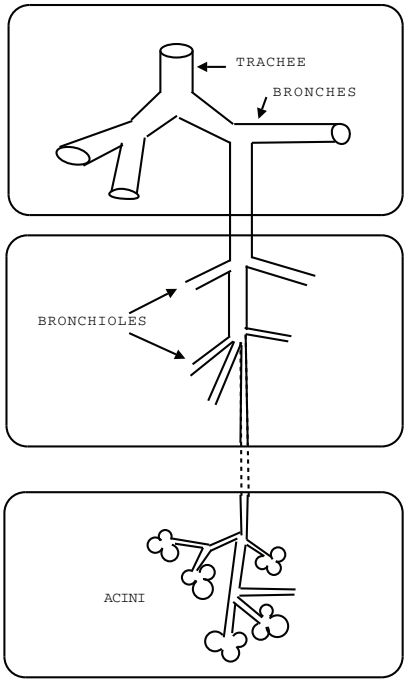


FIG. 2.9: Décomposition de l'arbre respiratoire en trois parties.

L'appareil respiratoire a une géométrie complexe : pour un calcul direct, c'est un arbre dichotomique composé de 23 générations qui devrait être maillé (voir Figure. 2.9). Ceci est infaisable numériquement.

Jusqu'à présent, les bronches distales situées à partir de la neuvième génération ne sont pas visualisables, ni segmentables par le matériel médical d'imagerie usuel. Afin de pallier les limites de l'imagerie et de mieux comprendre les différentes pneumopathies on comprend donc la nécessité d'élaborer un modèle simple mais réaliste.

Le modèle pour l'appareil respiratoire que l'on considère a déjà été décrit par GRANDMONT, MADAY et MAURY dans [GMM05] et on trouve également dans [BGM09, GMS08, SA07] des modèles similaires. Notons que le même type de modèle multi–échelles apparaît également pour la simulation de flux sanguins (voir par exemple [QRV01, QV03, VCFJT06]).

L'idée est de décomposer l'appareil respiratoire en trois parties :

- **la partie proximale** (jusqu'aux générations 7–9), où les équations de Navier–Stokes incompressibles sont vérifiées pour décrire le fluide ;
- **la partie distale** (de la 8–10e génération jusqu'à la 17e) où l'on peut supposer que la loi de Poiseuille est satisfaite dans chaque bronchiole ;
- **les acini**, où a lieu la diffusion de l'oxygène, nichés dans le parenchyme. Nous supposerons que la pression est uniforme dans les acini, égale à P_a (disons une moyenne de la pression alvéolaire). Le mouvement du diaphragme et du parenchyme est décrit par le modèle mécanique 0D détaillé précédemment (voir Section 2.4.1).

La Figure 2.10 illustre de manière synthétique la décomposition multi–échelles : dans la partie proximale Ω nous supposons que les équations de Navier–Stokes sont vérifiées et qu'elles sont couplées par le biais des conditions limites avec des flux de Poiseuille qui sont eux–mêmes couplés au modèle mécanique élastique. Le ressort décrit le mouvement du muscle diaphragmatique, supposé unidirectionnel.

Dans la partie supérieure Ω (voir Figure 2.10), nous supposons que les équations de Navier–Stokes sont satis-

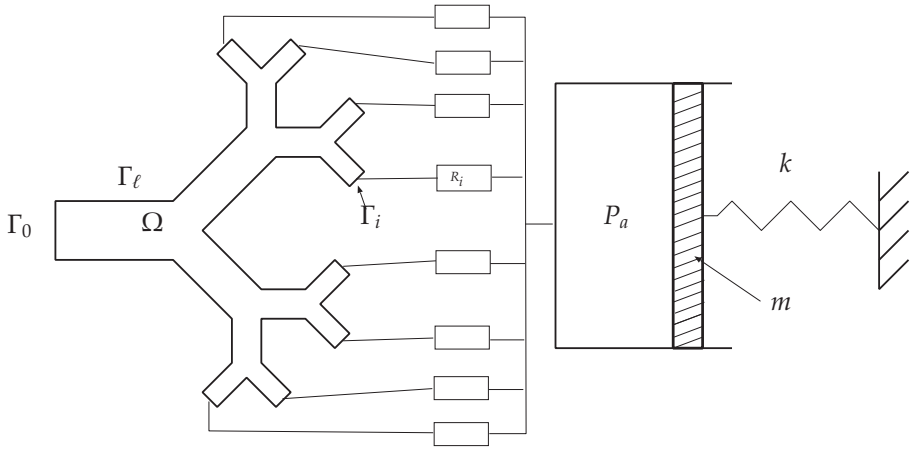


FIG. 2.10: Modèle multi–échelle.

faites :

$$\begin{cases} \rho\dfrac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \eta\triangle\mathbf{u} + \nabla p &= 0, & \text{sur } (0,T)\times\Omega, \\ \nabla\cdot\mathbf{u} &= 0, & \text{sur } (0,T)\times\Omega, \\ \mathbf{u} &= 0, & \text{sur } (0,T)\times\Gamma_\ell, \\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &= -\Pi_0\mathbf{n} & \text{sur } (0,T)\times\Gamma_0, \\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &= -\Pi_i\mathbf{n} & \text{sur } (0,T)\times\Gamma_i \quad i=1,\dots,N, \end{cases} \tag{2.13}$$

où $\mathbf{u}$ et p sont respectivement la vitesse du fluide et la pression du fluide. Sur la paroi latérale Γ_ℓ, nous imposons une condition de non–glissement sur la vitesse, tandis que sur les bords artificiels Γ_i, $0 \le i \le N$ nous considérons une force de pression exercée sur le bord. La condition de non–glissement sur les parois latérales est justifiée par la présence de mucus (voir page 75).

La pression Π_0 est donnée égale à P_0 la pression atmosphérique, tandis que les pressions Π_i sont inconnues et dépendent des parties en aval.

Pour chacun des sous–arbres de la partie distale, nous supposons l'arbre dichotomique et le flot laminaire. Ainsi, par analogie avec un réseau électrique (voir par exemple [BT06]), nous pouvons considérer que le flot est caractérisé par une unique résistance équivalente R_i pour l'ensemble du sous–arbre bronchique, dépendant de la résistance de chaque petite bronche (voir par exemple [Olu99]). Chacun des sous–arbres est alors remplacé par un domaine cylindrique, où le flot vérifie la loi de Poiseuille qui relie la différence de pression aux extrémités du tube : P_i (pression sur Γ_i) - P_a (pression alvéolaire) et le débit du fluide $\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}$:

$$\Pi_i - P_a = R_i \int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n},\ R_i \ge 0. \tag{2.14}$$

Grâce à la relation (2.14), les conditions aux bords aux sorties Γ_i s'écrivent :

$$\eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} = -P_a\mathbf{n} - R_i\left(\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n} \text{ sur } \Gamma_i, \quad i=1,\dots,N. \tag{2.15}$$

Ce sont des conditions non–standards et non–locales qui lient le tenseur des contraintes du fluide et son flux. Ceci apparaît également dans la modélisation des flux sanguins (voir [VCFJT06]). Il est à noter qu'elles créent de la dissipation dans le système (voir [BGM09]). Nous appellerons ce type de condition des conditions aux bords *naturellement dissipatives*.

Enfin, le poumon est enfermé dans une boîte connectée à un ressort qui dirige le mouvement du diaphragme et du parenchyme, comme présenté à la section précédente. Nous avons vu qu'alors l'équation satisfaite par la position x du diaphragme s'écrit :

$$m\ddot{x} = -kx + (RS^2 + \mu)\dot{x} + f_{ext} + f_P, \tag{2.16}$$

où dans ce cas les forces appliquées au piston sont la somme de f_{ext} la force musculaire et f_P la force de pression. f_{ext} est la force développée par le diaphragme durant l'inspiration et l'expiration. C'est la force qui dirige toute la respiration. Le couplage de cette simple équation différentielle à la partie supérieure du modèle, intervient via la définition de f_P qui correspond à la force de pression exercée par le fluide sur le muscle. Rappelons que R est la résistance des voies aériennes et μ la résistance des tissus. Ici comme la résistance des voies aériennes est déjà prise en compte dans le modèle fluide nous avons donc $R = 0$. Si S désigne la surface de la partie mobile de la boîte, correspondant à la surface du diaphragme, nous avons

$$f_P = P_a S, \tag{2.17}$$

d'où

$$P_a = \frac{m}{S}\ddot{x} - frac\mu S\dot{x} + \frac{k}{S}x - \frac{f_{ext}}{S}. \tag{2.18}$$

De plus, puisqu'on suppose ici le parenchyme constitué de matériau incompressible, la somme des flux aux sorties est égale à la variation de volume de la boîte modélisant le parenchyme, ainsi nous obtenons

$$S\dot{x} = \sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{u}\cdot\mathbf{n}. \tag{2.19}$$

Notons que puisque le fluide est incompressible et puisque $\mathbf{u} = 0$ sur Γ_ℓ, nous avons

$$\sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{u}\cdot\mathbf{n} = -\int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n}. \tag{2.20}$$

Finalement, le système couplé peut être réécrit comme suit

$$\left\{\begin{array}{rcll} \rho\dfrac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \eta\triangle\mathbf{u} + \nabla p &=& 0\,, & \text{sur } (0,T)\times\Omega\,,\\ \nabla\cdot\mathbf{u} &=& 0\,, & \text{sur } (0,T)\times\Omega\,,\\ \mathbf{u} &=& 0\,, & \text{sur } (0,T)\times\Gamma_\ell\,,\\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &=& -P_0\mathbf{n}\,, & \text{sur } (0,T)\times\Gamma_0\,,\\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &=& -P_a\mathbf{n} - R_i\left(\displaystyle\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n}\,, & \text{sur } (0,T)\times\Gamma_i\,,\ i=1\,,\ \ldots\,,\ N\,,\\ & & & \\ m\ddot{x} + \mu\dot{x} + kx &=& f_{ext} + SP_a\,, & \\ S\dot{x} &=& \displaystyle\sum_{i=1}^{N}\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n} = -\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}. & \end{array}\right. \tag{2.21}$$

Les inconnues sont la vitesse et la pression du fluide, $\mathbf{u}$ et p ainsi que la pression alvéolaire P_a et x le déplacement du thorax. On remarque que, connaissant P_a, on peut calculer $\mathbf{u}$, p et x, la pression alvéolaire solution correspondant à la pression telle que $S\dot{x} = -\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}$. Ce système doit être complété par des conditions initiales adaptées

$$\left\{\begin{array}{l} (\mathbf{u}, x, \dot{x})|_{t=0} = (\mathbf{u}_0, x_0, x_1), \text{ avec}\\ \nabla\cdot\mathbf{u}_0 = 0\,,\ \mathbf{u}_0 = 0 \text{ sur } \Gamma_\ell\,,\ \text{ et } Sx_1 = -\displaystyle\int_{\Gamma_0}\mathbf{u}_0\cdot\mathbf{n}. \end{array}\right. \tag{2.22}$$

Une particularité de ce système est que toutes les sorties $\Gamma_i, 1 \le i \le N$ sont couplées.

Analyse mathématique

L'analyse mathématique de ce système a été réalisée par BAFFICO, GRANDMONT et MAURY dans [BGM09]. Dans ce travail, les auteurs montrent l'existence de solutions régulières, dans un sens à préciser, localement en temps pour des données grandes et globalement en temps dans le cas particulier où les données sont petites et la constante de raideur du ressort nulle.

Au Chapitre 3, à la Section 3.2 nous discuterons des difficultés liées à l'analyse mathématique du problème (2.21)–(2.22) et nous préciserons un peu plus l'énoncé de ce théorème [BGM09].

Étude numérique

L'originalité de notre travail réside dans la méthode employée pour simuler numériquement le système (2.21)–(2.22).

Dans leurs travaux, BAFFICO, GRANDMONT et MAURY [BGM09] mettent en place un code permettant de traiter les conditions aux bords dissipatives. Il est important de noter que, puisque ces conditions relient tous les degrés de liberté des bords de sortie, les logiciels éléments finis commerciaux ne permettent pas de traiter facilement ces conditions dissipatives, ni le couplage avec le ressort. Essentiellement le traitement direct de ces conditions aux bords nécessite la modification du produit matrice–vecteur pour la résolution du système linéaire associé au problème, et donc demande de rentrer en profondeur dans le code.

La méthode que nous utiliserons permet de nous affranchir astucieusement de telles contraintes. L'idée est de calculer des solutions élémentaires, chacune solution d'un problème de Stokes où on impose une condition de Neumann égale à 1 sur une unique sortie et des conditions de Neumann homogènes sur toutes les autres. La stratégie est ensuite d'écrire la solution comme combinaison linéaire de solutions élémentaires pré–calculées et d'un terme correctif correspondant à la partie non–stationnaire du problème. Chacune de ces quantités est solution d'un problème de Stokes avec des conditions aux bords mixtes standards Dirichlet–Neumann. Les coefficients de la combinaison linéaire sont ensuite calculés de façon à ce que les conditions aux bords dissipatives soient vérifiées. L'avantage de cette méthode est qu'elle est facilement implémentable dans les logiciels éléments finis courants tels que `Freefem++` [HLHOP]. Néanmoins cette méthode reposant sur un principe de linéarité, il est nécessaire de traiter les termes de convection soit de façon explicite soit par la méthode des caractéristiques. La deuxième partie du Chapitre 3 détaille cette méthode et donne quelques simulations numériques.

2.4.3 Un modèle pour la diffusion de l'oxygène dans l'arbre bronchique

La deuxième partie de ce travail (Chapitre 4), concerne la modélisation des échanges gazeux dans le poumon. Dans cette section, nous donnons les idées essentielles sur lesquelles repose notre modèle ainsi que le système obtenu, annonçant la description détaillée effectuée au Chapitre 4.

Comme déjà mentionné, l'efficacité de l'échange gazeux au niveau de la membrane dépend de quatre facteurs :

- ⋆ la différence de pressions partielles en oxygènes entre l'alvéole et le sang pulmonaire ;
- ⋆ le ratio ventilation/perfusion qui mesure le rapport entre fréquence respiratoire et fréquence cardiaque ;
- ⋆ la perméabilité de la membrane ;
- ⋆ la surface d'échgange efficace.

Si les trois premiers facteurs sont soit fixés et donnés par la littérature (pression partielle en oxygène dans le sang, ratio ventilation/perfusion) soit facilement définissables (pression partielle en oxygène dans l'alvéole), le dernier correspond à une notion pas aisément quantifiable. En effet, cette donnée est fortement dépendante de la géométrie et du régime respiratoire auquel on se place. Dans [GH06], les auteurs définissent la notion de capacité de diffusion qui correspond au produit de la perméabilité avec la surface efficace. Puisqu'ils considèrent le poumon comme un tout, sans prendre en compte réellement la géométrie, ils donnent deux valeurs, l'une correspondant au régime de repos, l'autre à l'exercice. Une première partie de notre travail, consiste à retrouver les flux d'oxygène absorbé moyens à l'aide d'un modèle mon–compartiment.

Le modèle obtenu s'écrit de la façon suivante :

$$\begin{cases} \dot{x} &= u, \\ \dot{u} &= \dfrac{1}{m}\left(f_{ext} - \left(\dfrac{R_0 S^2}{1+\theta S x/V_B^0} + \mu\right)u - kx\right), \\ \dot{V}_{O_2} &= \beta^*\,\rho_{V/P}\left(V_A^0 + (1-\theta)Sx\right)^{2/3}\,(c_A\,P_0 - P_{\text{sang}}), \\ \dot{c}_A &= \dfrac{1}{V_A^0 + (1-\theta)Sx}\left((1-\theta)S\dot{x}\,[\dot{x}\geq 0\,?\,c_0 - c_A, 0] - \dot{V}_{O_2}\right), \end{cases}$$

où on utilise la notation suivante :

$$[y \geq 0\,?\,a, b] = \begin{cases} a & \text{si } y \geq 0, \\ b & \text{sinon.} \end{cases}$$

Les deux premières équations correspondent au modèle mécanique déjà présenté à la Section 2.4.1. La troisième équation compte le volume d'oxygène traversant la membrane ; nous reconnaissons $c_a P_0 - P_{\text{sang}}$ la différence des pressions partielles en oxygène de chaque côté de la membrane, $\rho_{V/P}$ le ratio ventilation/perfusion et $\beta^* \left(V_A^0 + (1-\theta)Sx\right)^{2/3}$ la capacité de diffusion. À l'aide des valeurs de [GH06], nous retrouvons les bons ordres de grandeur pour les volumes d'oxygène absorbé moyens.

Toutefois, ce modèle ne permet pas de rendre compte de l'hétérogénéité des échanges gazeux ; c'est pourquoi la deuxième partie de ce travail se tourne vers une description 1D du poumon, où la dimension d'espace correspond à la distance parcourue dans l'arbre depuis la bouche et sera notée ℓ. L'objectif de ce modèle est de montrer que grâce à la géométrie et à la mécanique, nous multiplions par 3 le volume d'oxygène absorbé, en passant d'un régime de respiration au repos à un régime de respiration à l'effort. Après avoir décrit la problématique et donné un aperçu des modèles déjà existants à la Section 4.3, nous proposons un nouveau modèle 1D à la Section 4.4. L'arbre bronchiques est alors modélisé par un tube de section variable. Nous supposons que ce tube représentant les bronches est connecté aux alvéoles qui sont séparés des vaisseaux sanguins par une membrane poreuse. Voir Figure 4.2 pour une représentation schématique. Le modèle complet obtenu est le suivant :

$$\dot{x} = u,$$

$$\dot{u} = \frac{1}{m}\left(f_{ext} - \left(\frac{R_0 S^2}{1 + \theta S x / V_B^0} + \mu\right) u - kx\right),$$

$$\partial_t s_b + \partial_\ell (s_b u^{\mathrm{f}}) = -\partial_t s_a,$$

$$\partial_t (s_b c_b) + \partial_\ell (s_b c_b u^{\mathrm{f}}) + \partial_\ell (-D_{O_2} s_b \partial_\ell c_b) = -\partial_t s_a [\dot{x} \geq 0\,?\,c_b, c_a] - W(\ell)\,(c_b - c_a),$$

$$\partial_t (s_a c_a) = -b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}) + \partial_t s_a [\dot{x} \geq 0\,?\,c_b, c_a] + W(\ell)\,(c_b - c_a).$$

À nouveau les deux premières lignes correspondent à la partie mécanique. La troisième équation traduit la conservation du volume d'air ; s_b désigne la section bronchique et s_a la section équivalente alvéolaire, de sorte que $s_a(\ell)\,\mathrm{d}\ell$ donne le volume alvéolaire compris entre ℓ et $\ell + \mathrm{d}\ell$, la vitesse du fluide est notée u^{f}. La quatrième équation correspond au bilan de masse pour l'oxygène écrit pour les bronches, tandis que la dernière équation correspond au bilan de masse écrit pour les alvéoles. c_a (resp. c_b) désigne la concentration en oxygène dans les alvéoles (resp. dans les bronches). Dans le membre de gauche de la quatrième équation, nous reconnaissons les termes de transport à la vitesse du fluide et de diffusion de l'oxygène.

Le terme commun aux deux dernières équations,

$$\partial_t s_a [\dot{x} \geq 0\,?\,c_b, c_a] + W(\ell)\,(c_b - c_a),$$

correspond aux termes d'échange entre bronche et alvéole ; le terme $\partial_t s_a [\dot{x} \geq 0\,?\,c_b, c_a]$ est un terme convectif correspondant au volume d'oxygène entrant dans l'alvéole dû à la dépression causée par la déformation mécanique de l'alvéole (à l'inspiration) ou sortant de l'alvéole dû à la surpression causée cette fois par la compression de l'alvéole (à l'expiration). Le terme $W(\ell)\,(c_b - c_a)$ est un terme de diffusion dans la direction 'orthogonale' entre l'alvéole et la bronche.

Le terme de perte dans la troisième équation,

$$-b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}),$$

correspond à la quantité d'oxygène absorbée qui passe dans le sang avec b la perméabilité, $\rho_{V/P}$ le ratio ventilation/perfusion, ℓ_a la densité de surface alvéolaire et $(P_0 c_a - P_{sang})$ la différence de pressions partielles en oxygène entre l'alvéole et le sang.

Nous terminons par des simulations numériques du système. Nous soulignons que ce travail est encore un travail en cours.

Chapitre 3

Une méthode numérique pour la simulation du flux d'air dans les poumons

Dans ce chapitre, nous nous intéressons à la simulation du flux d'air dans l'arbre bronchique. Le modèle que nous étudions, décrit à la Section 2.4.2, consiste en une description selon trois parties de l'arbre respiratoire. Ce modèle nous conduit, après discrétisation en temps, à un problème de Stokes avec des conditions aux bords dissipatives non–usuelles qui ne peuvent être implémentées facilement et directement dans la plupart des logiciels utilisant la méthode des éléments finis, en particulier `Freefem++` [HLHOP]. L'objectif ici, est d'apporter une méthode de résolution implémentable dans tout logiciel de calcul par éléments finis (EFM). Après une description de la méthode, nous l'illustrons par des simulations 2D. Ce travail, initié au CEMRACS'08 est le fruit d'une collaboration avec CELINE GRANDMONT, BÉRÉNICE GREC, BERTRAND MAURY et DRISS YAKOUBI. Cette méthode s'adapte très bien au cas de la dimension $d = 3$; l'implémentation 3D a été réalisée par DRISS YAKOUBI. L'essentiel des résultats est paru dans les *Proceedings* du CEMRACS sous le titre *Numerical Method for the 2D Simulation of the Respiration* [DGG+09].

3.1 Introduction

Nous nous intéressons au système suivant dérivé du modèle multi–échelles que nous avons décrit précédemment à la Section 2.4.2, page 85. Rappelons ci–dessous le système :

$$
\left\{
\begin{array}{rcll}
\rho\dfrac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \eta\triangle\mathbf{u} + \nabla p &=& 0\,, & \text{sur } (0,T)\times\Omega\,,\\
\nabla\cdot\mathbf{u} &=& 0\,, & \text{sur } (0,T)\times\Omega\,,\\
\mathbf{u} &=& 0\,, & \text{sur } (0,T)\times\Gamma_\ell\,,\\
\eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &=& -P_0\mathbf{n}\,, & \text{sur } (0,T)\times\Gamma_0\,,\\
\eta\nabla\mathbf{u}\cdot\mathbf{n} - p\mathbf{n} &=& -P_a\mathbf{n} - R_i\left(\displaystyle\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n}\,, & \text{sur } (0,T)\times\Gamma_i\,,\ i=1\,,\ \ldots\,,\ N\,,\\
\\
m\ddot{x} + \mu\dot{x} + kx &=& f_{ext} + SP_a\,,\\
S\dot{x} &=& \displaystyle\sum_{i=1}^{N}\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n} = -\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}.
\end{array}
\right. \tag{3.1}
$$

Ce système doit être complété par des conditions initiales adaptées

$$\begin{cases} (\mathbf{u}, x, \dot{x})|_{t=0} = (\mathbf{u}_0, x_0, x_1), \text{ avec} \\ \nabla \cdot \mathbf{u}_0 = 0\,,\ \mathbf{u}_0 = 0 \text{ sur } \Gamma_\ell\,, \quad \text{et } Sx_1 = -\displaystyle\int_{\Gamma_0} \mathbf{u}_0 \cdot \mathbf{n}. \end{cases} \tag{3.2}$$

Une particularité de ce système est que toutes les sorties $\Gamma_i, 1 \leq i \leq N$, sont couplées et que des termes non locaux apparaissent au bord.

Remarque 5 *En posant* $\overline{p} = p - P_a$*, en utilisant* (2.18) *et la préservation du volume total,* i.e $S\dot{x} = -\displaystyle\int_{\Gamma_0} \mathbf{u} \cdot \mathbf{n}$*, le système couplé peut se simplifier comme suit*

$$\begin{cases} \rho\dfrac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \eta\triangle\mathbf{u} + \nabla\overline{p} &=& 0\,, & \textit{sur } (0,T)\times\Omega\,, \\ \nabla\cdot\mathbf{u} &=& 0\,, & \textit{sur } (0,T)\times\Omega\,, \\ \mathbf{u} &=& 0\,, & \textit{sur } (0,T)\times\Gamma_\ell\,, \\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - \overline{p}\mathbf{n} &=& -P_0\mathbf{n} - \dfrac{f_{ext}}{S}\mathbf{n} - \dfrac{m}{S^2}\dfrac{d}{dt}\left(\displaystyle\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n} & \\ & & -\dfrac{\mu}{S^2}\left(\displaystyle\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n} - \dfrac{k}{S^2}\left(\displaystyle\int_0^t\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n} - Sx_0\right)\mathbf{n}\,, & \textit{sur } (0,T)\times\Gamma_0\,, \\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - \overline{p}\mathbf{n} &=& -R_i\left(\displaystyle\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\mathbf{n}\,, & \textit{sur } (0,T)\times\Gamma_i. \end{cases} \tag{3.3}$$

En conséquence, le système (2.21) *se réduit aux équations de Navier–Stokes où la pression du fluide est remplacée par la différence entre la pression du fluide et la pression alvéolaire, munies des conditions aux bords* naturellement dissipatives*. Le lecteur notera que l'hypothèse d'incompressibilité est essentielle ici.*

Remarque 6 *Nous avons utilisé la relation* $\displaystyle\sum_{i=1}^{N}\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n} = -\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}$ *afin de simplifier le système* (2.21)*, découpler les sorties* Γ_i *et obtenir* (3.3)*. Néanmoins cette hypothèse n'est pas nécessaire dans la suite.*

Comme nous l'avons mentionné ce système a été étudié mathématiquement dans [BGM09]. Dans une première partie de ce chapitre nous nous attacherons à décrire les difficultés liées à l'étude mathématique de ce système notamment et nous énoncerons le résultat d'existence et l'unicité de solution en temps court ainsi que l'existence globale dans un cas particulier dans [BGM09]. En particulier, nous noterons la difficulté d'obtenir des estimations d'énergie pour le terme non–linéaire à savoir le flux d'énergie cinétique aux bords Γ_i. Toutefois, les auteurs obtiennent des estimations d'énergie en dimension $d = 2$, localement en temps pour de petites données. Afin d'obtenir l'existence de solution sans restriction sur les données aussi bien en dimension 2 ou 3 les auteurs travaillent avec des solutions plus régulières. Par une approche de type Galerkin, ils montrent qu'il existe une unique solution localement en temps. Sous des hypothèses supplémentaires ($k = 0$, données petites) ils obtiennent l'existence globale.

Dans une deuxième partie, nous décrirons la méthode numérique que nous avons développée et implémentée. Après une discrétisation en temps du problème couplé (basée sur la méthode des caractéristiques pour le terme de convection), nous obtenons un système de Stokes muni de conditions aux bords non–standards mettant en jeu les flux d'air aux interfaces. La difficulté consiste alors à traiter ces conditions aux bords dissipatives à l'aide d'un logiciel éléments finis par exemple `FreeFEM++`. La stratégie choisie ici est d'écrire la solution comme combinaison linéaire de solutions élémentaires pré–calculées et d'un terme correctif correspondant à la partie non–stationnaire du problème. Chacune de ces quantités est solution d'un problème de Stokes avec des conditions aux bords mixtes standards Dirichlet–Neumann. Les coefficients de la combinaison linéaire sont ensuite calculés de façon à ce que les conditions aux bords dissipatives soient satisfaites. Le lecteur doit noter qu'une telle méthode a déjà été décrite dans un contexte légèrement différent (voir [FGNQ02, VV07]). Dans ces papiers, les auteurs ont a traiter des conditions aux bords défectives[7]. Afin de calculer proprement la solution, ils voient leurs conditions aux

[7]Dans ces travaux, les auteurs imposent la pression moyenne ou le flux moyen aux interfaces Γ_i, $1 \leq i \leq N$. Ces conditions ne permettent pas de fermer le système et sont donc appelées défectives.

bords défectives en tant que contraintes et obtiennent ainsi une nouvelle formulation à l'aide de multiplicateurs de Lagrange. Numériquement, cette formulation revêt des aspects proches des nôtres et afin de réduire leur problème ils utilisent une méthode de *spliting*. Enfin, dans la dernière partie nous présenterons les résultats numériques, et vérifierons que pour un certain jeu de paramètres, notre modèle reproduit certains aspects de la respiration.

3.2 Analyse mathématique

3.2.1 Formulation variationnelle

On introduit l'espace fonctionnel suivant

$$V = \{\mathbf{v} \in H^1(\Omega)^d\,,\ \nabla\cdot\mathbf{v} = 0, \mathbf{v} = 0 \ \text{ sur } \Gamma_\ell\}. \tag{3.4}$$

Supposons que toutes les inconnues soient suffisamment régulières, nous multiplions les équations de Navier–Stokes par des fonctions tests $\mathbf{v}$ s'annulant sur Γ_ℓ, et l'équation mécanique par $-(1/S)\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}$. À l'aide de

$$x = x_0 - \frac{1}{S}\int_0^t\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n},$$

nous obtenons

$$\begin{aligned}
&\rho\int_\Omega \partial_t\mathbf{u}\cdot\mathbf{v} + \rho\int_\Omega(\mathbf{u}\cdot\nabla)\mathbf{u}\cdot\mathbf{v} + \eta\int_\Omega\nabla\mathbf{u}:\nabla\mathbf{v} + \sum_{i=1}^N R_i\left(\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_i}\mathbf{v}\cdot\mathbf{n}\right)\\
&+\frac{m}{S^2}\left(\int_{\Gamma_0}\partial_t\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right) + \frac{k}{S^2}\left(\int_0^t\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right) + \frac{\mu}{S^2}\left(\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)\\
&-\int_\Omega p\nabla\cdot\mathbf{v} + P_a\left(\sum_{i=1}^N\int_{\Gamma_i}\mathbf{v}\cdot\mathbf{n} + \int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)\\
&= -P_0\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n} - \frac{f_{ext}}{S}\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n} + \frac{k}{S^2}Sx_0\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right), \quad \forall\mathbf{v}\in H^1(\Omega)^d \text{ avec } \mathbf{v}|_{\Gamma_\ell} = 0.
\end{aligned}$$

Ensuite, en considérant des fonctions tests $\mathbf{v}$ à divergence nulle (*i.e* $\mathbf{v}\in V$), nous obtenons une seconde formulation variationnelle du problème couplé (3.3)

$$\begin{aligned}
&\rho\int_\Omega \partial_t\mathbf{u}\cdot\mathbf{v} + \rho\int_\Omega(\mathbf{u}\cdot\nabla)\mathbf{u}\cdot\mathbf{v} + \eta\int_\Omega\nabla\mathbf{u}:\nabla\mathbf{v} + \sum_{i=1}^N R_i\left(\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_i}\mathbf{v}\cdot\mathbf{n}\right)\\
&+\frac{m}{S^2}\left(\int_{\Gamma_0}\partial_t\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right) + \frac{k}{S^2}\left(\int_0^t\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right) + \frac{\mu}{S^2}\left(\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)\\
&= -P_0\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n} - \frac{f_{ext}}{S}\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n} + \frac{k}{S}x_0\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right), \quad \forall\mathbf{v}\in V.
\end{aligned} \tag{3.5}$$

Remarquons qu'ici nous avons exprimé toutes les quantités à l'aide de la vitesse du fluide. La vitesse du ressort peut être retrouvée facilement à l'aide de la relation $S\dot{x} = -\int_{\Gamma_0}\mathbf{u}\cdot\mathbf{n}$.

3.2.2 Bilan d'énergie

Formellement, nous pouvons écrire un bilan d'énergie pour le système couplé. Pour ce faire, on choisit $\mathbf{v} = \mathbf{u}$ comme fonction test dans (3.5). En réécrivant par intégration par parties le terme convectif comme un flux d'énergie

cinétique à l'entrée et aux sorties et en utilisant le fait que $S\dot{x} = -\int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n}$, on obtient :

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\rho}{2}\int_\Omega |\mathbf{u}|^2 + \frac{m}{2}|\dot{x}|^2 + \frac{k}{2}|x|^2\right) + \eta\int_\Omega |\nabla\mathbf{u}|^2 + \sum_{i=1}^{N} R_i\left(\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)^2 + \mu|\dot{x}|^2 = -\frac{\rho}{2}\sum_{i=0}^{N}\int_{\Gamma_i}|\mathbf{u}|^2(\mathbf{u}\cdot\mathbf{n}) + P_0 S\dot{x} + f_{ext}\dot{x}, \quad (3.6)$$

où :

- $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\rho}{2}\int_\Omega |\mathbf{u}|^2 + \frac{m}{2}|\dot{x}|^2 + \frac{k}{2}|x|^2\right)$ correspond à la variation d'énergie totale du système, les deux premiers termes représentant la variation totale d'énergie cinétique, le dernier la variation de l'énergie mécanique du ressort ;
- $\eta\int_\Omega |\nabla\mathbf{u}|^2$ correspond à la dissipation d'énergie due à la viscosité du fluide ;
- $\sum_{i=1}^{N} R_i\left(\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\right)^2$ est l'énergie dissipée engendrée par la résistance des voies aériennes dans les sous-arbres ;
- $\mu|\dot{x}|^2$ est l'énergie dissipée par les tissus pulmonaires ;
- $-\frac{\rho}{2}\sum_{i=0}^{N}\int_{\Gamma_i}|\mathbf{u}|^2(\mathbf{u}\cdot\mathbf{n})$ est le flux d'énergie cinétique à travers l'entrée et les sorties ;
- $P_0 S\dot{x}$ est la puissance développée par la pression atmosphérique à l'entrée ;
- $f_{ext}\dot{x}$ est la puissance développée par les forces extérieures exercées sur le diaphragme.

Ce bilan d'énergie permettrait d'obtenir des estimations d'énergie, mais pour cela il faudrait estimer correctement les flux d'énergie cinétique aux interfaces à l'aide de l'énergie cinétique totale et de l'énergie dissipée. Cependant on peut obtenir un contrôle du flux d'énergie cinétique localement en temps, pour des données petites mais uniquement en dimension $d = 2$. Cette estimation permettrait de montrer l'existence de solutions faibles (dans les espaces d'énergie). Afin de traiter le cas de la dimension 2 et 3, sans restriction sur les données dans [BGM09] l'existence de solutions plus régulières est considéré.

Remarque 7 *Dans [GMM05], les auteurs contournent cette difficulté en choisissant d'autres conditions limites ; leurs conditions aux bords prennent en compte le tenseur des contraintes total, à savoir la pression p est remplacée par la pression totale* $p + \rho\frac{|u|^2}{2}$ *faisant ainsi disparaître le flux d'énergie cinétique aux bords du bilan d'énergie. L'estimation d'énergie étant insufisante pour traiter le terme non linéaire, ils passent ensuite en Fourier en temps pour obtenir de la compacité en temps et passer à la limite.*

Dans [GMS08], une autre stratégie est choisie : les auteurs supposent que sur chaque sortie la trace de la vitesse du fluide est proportionnelle à un profil donné $\mathbf{U}_i$*, soit* $\mathbf{u} = \lambda_i\mathbf{U}_i$ *sur* Γ_i*. Il s'agit de conditions plus restrictives, mais en se limitant à un seul degré de liberté en vitesse pour l'entrée et chacune des sorties du domaine de calcul, on peut contrôler le flux d'énergie cinétique sur les frontières artificielles.*

3.2.3 Réécriture du problème

Dans cette partie, nous réécrivons la formulation variationnelle (3.5) sous une forme abstraite.
Soient Ω un domaine borné de $\mathbb{R}^d$, $d = 2, 3$, Γ_0 l'interface d'entrée, Γ_i les sorties (voir Figure 2.10). Notons Γ_ℓ la paroi latérale. Soit l'espace fonctionnel suivant :

$$H = \overline{V}^{L^2},$$

où V est défini en (3.4). Nous noterons $(\cdot,\cdot)_0$ le produit scalaire sur $H \times H$ défini par

$$(\mathbf{v},\mathbf{w})_0 = \rho \int_\Omega \mathbf{v}\cdot\mathbf{w} + \frac{m}{S^2}\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{w}\cdot\mathbf{n}\right),$$

et $\|\cdot\|_0$ la norme associée. Ensuite, posons

$$a_1(\mathbf{v},\mathbf{w}) = \eta\int_\Omega \nabla\mathbf{v}:\nabla\mathbf{w} + \sum_{i=1}^{N} R_i\left(\int_{\Gamma_i}\mathbf{v}\cdot\mathbf{n}\right)\left(\int_{\Gamma_i}\mathbf{w}\cdot\mathbf{n}\right) + \frac{\mu}{S^2}\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{w}\cdot\mathbf{n}\right)$$

et

$$b(\mathbf{u},\mathbf{v},\mathbf{w}) = \rho\int_\Omega (\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}$$

qui sont respectivement les formes bilinéaire et trilinéaire sur V associées au problème. Enfin notons

$$a_2(\mathbf{v},\mathbf{w}) = \frac{k}{S^2}\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)\left(\int_{\Gamma_0}\mathbf{w}\cdot\mathbf{n}\right).$$

Finalement, pour compléter le formalisme, introduisons l'opérateur A dont les vecteurs propres constitueront par la suite la base de Galerkin utilisée pour construire une suite de solutions approchées.

Définition 3.2.1 (Opérateur A) *L'opérateur A est défini sur H comme suit :*

$$\begin{aligned} D(A) &= \{\mathbf{v}\in V,\ |a_1(\mathbf{v},\mathbf{w})| \le C\|\mathbf{v}\|_0,\ \forall \mathbf{w}\in V\},\\ (A\mathbf{v},\mathbf{w})_0 &= a_1(\mathbf{v},\mathbf{w}),\ \forall\,(\mathbf{v},\mathbf{w})\in D(A)\times V. \end{aligned} \tag{3.7}$$

On donne maintenant la définition de solution du système (3.1)–(3.2).

Définition 3.2.2 *On dit que* $\mathbf{u}$ *est solution du problème* (3.1)–(3.2) *sur* $[0,T]$ *si*

$$\begin{cases} \mathbf{u}\in L^2(0,T;D(A))\cap L^\infty(0,T;V)\cap H^1(0,T;H),\\ \dfrac{\mathrm{d}}{\mathrm{d}t}(\mathbf{u},\mathbf{v})_0 + a_1(\mathbf{u},\mathbf{v}) + b(\mathbf{u},\mathbf{u},\mathbf{v}) + a_2\left(\displaystyle\int_0^t \mathbf{u},\mathbf{v}\right) = \ell(\mathbf{v}),\ \ \forall\mathbf{v}\in V, \end{cases} \tag{3.8}$$

avec

$$\ell(\mathbf{v}) = -P_0\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n} - \frac{f_{ext}}{S}\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n} + \frac{k}{S^2}Sx_0\left(\int_{\Gamma_0}\mathbf{v}\cdot\mathbf{n}\right)$$

et la condition initiale

$$\mathbf{u}(0) = \mathbf{u}_0.$$

3.2.4 Existence et unicité d'une solution

Dans [BGM09], les auteurs donnent deux types de résultats d'existence : l'un local en temps, pour de grandes données et l'autre global en temps pour des données suffisamment petites dans le cas particulier où la raideur du ressort est nulle. Nous redonnons ici leur résultat d'existence. Tout d'abord explicitons deux hypothèses utiles :

Hypothèse 1 *Ω est un domaine de $\mathbb{R}^d$ est régulier par morceaux dans le sens où chaque composante de la frontière Γ_i, $i = 1,\dots,N$, et Γ_ℓ est régulière. De plus, on suppose que l'angle entre les Γ_i et la paroi latérale Γ_ℓ est égal à $\pi/2$.*

Hypothèse 2 *Pour $T > 0$, supposons*

$$P_0\in L^2(0,T),\quad \mathbf{u}_0\in V,\quad \textit{et}\quad R_i\ge 0,\ i=1,\dots,N$$

Théorème 3.2.1 *Sous les Hypothèses 1 et 2, il existe un intervalle de temps* $[0, t^*]$ *sur lequel le système* (3.1)–(3.2) *admet une unique solution dans le sens explicité à la Définition 3.2.2. De plus, dans le cas particulier où* $k = 0$ *et pour des données petites (condition initiale et force exercée suffisamment petite en norme* L^∞ *en temps), la solution peut être définie globalement en temps et il existe une constante* C *telle que pour tout* $t > 0$,

$$\eta \|\nabla \mathbf{u}\|^2_{L^2(\Omega)} + \sum_{i=1}^{N} R_i \left(\int_{\Gamma_i} \mathbf{u} \cdot \mathbf{n} \right)^2 + \frac{\mu}{S^2} \left(\int_{\Gamma_0} \mathbf{u} \cdot \mathbf{n} \right)^2 \leq C.$$

Dans leur preuve BAFFICO, GRANDMONT et MAURY [BGM09] suivent la méthode utilisée dans [RHT96] : ils construisent une suite de solutions approchées à l'aide d'une méthode de Galerkin (les vecteurs de la base de Galerkin étant les vecteurs propres associés à l'opérateur A), puis établissent des bornes uniformes qui leur permettent de passer à la limite montrant ainsi l'existence d'au moins une solution de (3.8). Ils montrent finalement qu'une telle solution est unique. La restriction au cas $k = 0$ résulte de l'incapacité à contrôler le déplacement du ressort. Cependant, si on suppose *a priori* que celui-ci est borné par une constante suffisamment petite, la méthode permet d'obtenir le résultat d'existence globale pour $k \neq 0$.

Remarque 8 *Dans [RHT96], les auteurs montrent l'existence d'une unique solution régulière au système de Navier–Stokes lorsqu'on impose la pression moyenne ou bien le flux moyen en entrée et en sortie. Ils utilisent la méthode dite du 'do nothing'. Par cette technique, une formulation faible particulière est donnée permettant de satisfaire aux conditions de pression moyenne ou de flux moyen selon le cas. Cette formulation contient de plus implicitement des conditions aux bords (de type Neumann) qui sélectionnent une solution particulière parmi toutes les solutions admissibles physiquement pour le problème originel. Dans [QV03], les auteurs, dans le cadre de la simulation des écoulements sanguins, s'intéressent au système de Navier–Stokes avec les conditions aux bords suivantes :* $\eta \nabla \mathbf{u} \cdot \mathbf{n} - p\mathbf{n} + K_i(\mathbf{u} \cdot \mathbf{n})\mathbf{n} = -p_i \mathbf{n}$ *où* $p_i = p_i(t)$ *est considéré comme une fonction donnée ne dépendant que du temps. Ils montrent en particulier l'existence d'une solution au système de Navier–Stokes couplé à un modèle 0D par une méthode de point fixe. Dans notre modèle, la nature du couplage ne permet pas d'appliquer facilement les mêmes arguments. De plus ils obtiennent l'existence de solution si* K_i *(qui est un analogue des résistances* R_i*) est suffisament grand. Dans [BGM09], les auteurs s'affranchissent de cette contrainte par le choix de l'opérateur A, qui prend alors en compte la partie dissipative provenant de la résistance des sous–arbres.*

3.3 Méthode numérique

Dans cette partie nous présentons la méthode numérique retenue pour la simulation du système couplé Navier–Stokes/masse–ressort.

Les logiciels éléments finis classiques ne permettent de traiter facilement ni les conditions aux bords dissipatives qui apparaissent dans notre système, ni le couplage avec le modèle élastique.

Une première stratégie possible [BGM09] consiste à modifier le produit matrice–vecteur pour la résolution du système linéaire associé au problème. Comme nous l'avons mentionné plus tôt, les conditions dissipatives conséquence des résistances des sous–arbres et du couplage avec le système masse–ressort modifient les formes bilinéaires dans la formulation variationnelle et par conséquent modifient la matrice EFM associée aux degrés de libertés pour la vitesse. Toutefois, cette modification peut être vue comme une perturbation additive de la matrice EFM originale,

$$A = A^\Omega + A^{res},$$

où A^Ω désigne la matrice EFM originale sur le domaine Ω et A^{res} la matrice associée aux résistances. À chaque étape, une méthode itérative est employée nécessitant le calcul du produit matrice–vecteur

$$Ax = A^\Omega x + A^{res} x.$$

Plutôt que d'assembler à chaque fois la matrice A, les deux termes sont calculés séparément. De plus, pour calculer le produit $A^{res}x$, la matrice A^{res} n'a pas besoin d'être assemblée puisque ce produit peut être écrit comme un produit tensoriel entre deux vecteurs, les résistances et une résistance équivalente en entrée qui prend en compte le couplage avec le ressort. Cette stratégie nécessite de rentrer en profondeur dans le code, ce que nous nous sommes interdit ici.

Une seconde stratégie, celle que nous avons retenue, consiste à décomposer la solution en une combinaison linéaire de solutions élémentaires pré–calculées et d'un terme correctif prenant en compte l'évolution temporelle, chacune de ces quantités étant solution d'un problème de Stokes avec conditions aux bords standard mixte Dirichlet–Neumann. Les coefficients de la combinaison linéaire sont calculés de façon à ce que les conditions aux bords dissipatives soient vérifiées.

Cette approche apparaît déjà dans [FGNQ02]. Dans ces travaux (voir aussi [VV07]), les auteurs proposent une méthode numérique pour traiter des conditions aux bords défectives. La méthode déjà proposée par [RHT96] pour traiter ce genre de conditions (voir Remarque 8) propose une formulation facile à discrétiser dans le cas du problème de pression moyenne. Toutefois, concernant le problème de flux moyen, la définition de l'espace fonctionnel rend l'implémentation moins facile. Partant de ce constat FORMAGGIA *et al* [FGNQ02] proposent une nouvelle approche basée sur les multiplicateurs de Lagrange.

Leur but est de résoudre le système de Navier–Stokes suivant :

$$\left\{\begin{array}{rl} \rho\partial_t\mathbf{u} + \rho\mathbf{u}\cdot\nabla\mathbf{u} - \eta\Delta\mathbf{u} + \nabla p = \mathbf{f}, & \text{sur } (0,T)\times\Omega, \\ \mathrm{div}\mathbf{u} = 0, & \text{sur } (0,T)\times\Omega, \\ \mathbf{u} = 0, & \text{sur } (0,T)\times\Gamma_\ell, \\ \langle\phi_i,\mathbf{u}\rangle = \displaystyle\int_{\Gamma_i}\mathbf{u}\cdot\mathbf{n}\,\mathrm{d}s = Q_i, & \text{pour } i = 1,\dots,N, \\ \eta\nabla\mathbf{u}\cdot\mathbf{n} - p\,\mathbf{n} = 0, & \text{sur } (0,T)\times\Gamma_0, \end{array}\right.$$

avec la condition initiale $\mathbf{u} = \mathbf{u}_0$ à $t = 0$. La difficulté de ce système est que les conditions aux bords sont défective ; en effet, on impose uniquement le flux moyen sur chaque sortie, ce qui ne permet pas de fermer le système. À cette fin, les auteurs [FGNQ02] proposent de résoudre le problème suivant :
Trouver $(\mathbf{u}, p, \lambda_1, \dots, \lambda_N) \in \overline{V}\times L^2(\Omega)\times\mathbb{R}^N$, tels que $\forall\mathbf{v}\in\overline{V}$, $\forall q\in L^2(\Omega)$

$$\left\{\begin{array}{rcll} \rho\left(\partial_t\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}, v\right) + \eta\left(\nabla\mathbf{u},\nabla\mathbf{v}\right) + \displaystyle\sum_{i=1}^{N}\lambda_i\langle\phi_i,\mathbf{v}\rangle - \left(p,\mathrm{div}\mathbf{v}\right) &=& \left(\mathbf{f},\mathbf{v}\right), & \\ \left(q,\mathrm{div}\mathbf{u}\right) &=& 0, & \\ \langle\phi_i,\mathbf{u}\rangle &=& Q_i, & i = 1,\dots,N \\ \mathbf{u} &=& \mathbf{u}_0, & \text{en } t = 0, \end{array}\right.$$

où $\overline{V} = \left\{\mathbf{v}\in H^1(\Omega)^d,\ \mathbf{v}|_{\Gamma_\ell} = 0\right\}$.

Dans un souci de clarté, nous présentons leur méthode en nous restreignant au problème de Stokes. Cela revient à trouver $\mathbf{u}\in\overline{V}$, $p\in L^2(\Omega)$ et $(\lambda_1,\dots,\lambda_N)\in\mathbb{R}^N$ tels que $\forall\mathbf{v}\in\overline{V}$, et $\forall q\in L^2(\Omega)$,

$$\left\{\begin{array}{rcll} \eta\left(\nabla\mathbf{u},\nabla\mathbf{v}\right) + \sum_{i=1}^{N}\lambda_i\langle\phi_i,\mathbf{v}\rangle - \left(p,\mathrm{div}v\right) &=& \left(\mathbf{f},\mathbf{v}\right) & \\ \left(q,\mathrm{div}\mathbf{u}\right) &=& 0, & \\ \langle\phi_i,\mathbf{u}\rangle &=& Q_i, & i = 1,\dots,N. \end{array}\right. \tag{3.9}$$

La solution est alors écrite comme combinaison linéaire de solutions de problèmes plus simples ; soit $(\tilde{\mathbf{u}},\tilde{p})\in\overline{V}\times L^2(\Omega)$ la solution de :

$$\forall(\mathbf{v},q)\in\overline{V}\times L^2(\Omega),\ \left\{\begin{array}{rcll} \eta\left(\nabla\tilde{\mathbf{u}},\nabla\mathbf{v}\right) - \left(\tilde{p},\mathrm{div}\mathbf{v}\right) &=& \left(\mathbf{f},\mathbf{v}\right), & \\ \left(q,\mathrm{div}\tilde{\mathbf{u}}\right) &=& 0, & \\ \tilde{\mathbf{u}} &=& 0, & \text{sur } \Gamma_\ell \\ \eta\nabla\tilde{\mathbf{u}}\cdot\mathbf{n} - \tilde{p}\mathbf{n} &=& 0, & \text{sur } \Gamma_0,\dots,\Gamma_N. \end{array}\right. \tag{3.10}$$

De plus, notons $(\mathbf{u}_i, p_i) \in \overline{V} \times L^2(\Omega)$ la solution du problème

$$\forall (\mathbf{v}, q) \in \overline{V} \times L^2(\Omega), \begin{cases} \eta\,(\nabla \mathbf{u}_i, \nabla \mathbf{v}) - (p_i, \mathrm{div}\mathbf{v}) &=& -\langle \phi_i, v\rangle, & \\ (q, \mathrm{div}\mathbf{u}_i) &=& 0, & \\ \mathbf{u}_i &=& 0, & \text{sur } \Gamma_\ell \\ \eta \nabla \mathbf{u}_i \cdot \mathbf{n} - p_i \mathbf{n} &=& 0, & \text{sur } \Gamma_j,\ j \neq i. \end{cases} \tag{3.11}$$

On pose ensuite

$$\mathbf{u} = \tilde{\mathbf{u}} + \sum_{i=1}^{N} \lambda_i \mathbf{u}_i,$$
$$\text{et} \quad p = \tilde{p} + \sum_{i=1}^{N} \lambda_i p_i.$$

Les λ_i sont déterminés de telle sorte que $\mathbf{u}$ vérifie la troisième équation de (3.9), soit

$$\langle \phi_i, \tilde{\mathbf{u}} \rangle + \sum_{j=1}^{N} \lambda_j \langle \phi_i, \mathbf{u}_j \rangle = Q_i, \quad i = 1, \ldots, N,$$

ce qui se réécrit sous la forme d'un système linéaire inversible.

Il est à noter que cette approche a l'avantage de pouvoir être implémentée sur un solveur Navier–Stokes classique avec conditions aux bords usuelles (Neumann, Dirichlet). Comme nous l'avons décrit précédemment, la méthode de résolution proposée réduit le problème à la résolution de $N + 1$ problèmes de Stokes non contraints (3.10)–(3.11). De plus, remarquons que pour une géométrie donnée, les N solutions $\mathbf{u}_i$ des problèmes (3.11) peuvent être calculées une fois pour toute.

Bien que nous soyons dans un cadre différent (ici pas de conditions défectives, notre système est entièrement déterminé), ces idées s'adaptent facilement à notre problème.

Dans cette section, nous allons présenter tout d'abord la méthode numérique sur un problème linéaire obtenu simplement en omettant les termes de convection (problème de Stokes). Nous expliquerons ensuite comment l'adapter au cas plus général.

3.3.1 Le système de Stokes

Dans cette section nous nous intéressons au système linéaire couplé suivant :

$$\begin{cases} \rho \dfrac{\partial \mathbf{u}}{\partial t} - \eta \Delta \mathbf{u} + \nabla p &=& 0, & \text{sur } (0,T) \times \Omega\,, \\ \nabla \cdot \mathbf{u} &=& 0, & \text{sur } (0,T) \times \Omega, \\ \mathbf{u} &=& 0, & \text{sur } (0,T) \times \Gamma_\ell, \\ \eta \nabla \mathbf{u} \cdot \mathbf{n} - p\mathbf{n} &=& -P_0 \mathbf{n} - \dfrac{f_{ext}}{S}\mathbf{n} - \dfrac{m}{S^2}\dfrac{d}{dt}\left(\displaystyle\int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n}\right)\mathbf{n} & \\ & & -\dfrac{k}{S^2}\left(\displaystyle\int_0^t \int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n} - Sx_0\right)\mathbf{n} - \dfrac{\mu}{S^2}\left(\displaystyle\int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n}\right)\mathbf{n}, & \text{sur } (0,T) \times \Gamma_0, \\ \eta \nabla \mathbf{u} \cdot \mathbf{n} - p\mathbf{n} &=& -R_i\left(\displaystyle\int_{\Gamma_i} \mathbf{u}\cdot\mathbf{n}\right)\mathbf{n}\,, & \text{sur } (0,T) \times \Gamma_i. \end{cases} \tag{3.12}$$

On commence par discrétiser en temps le problème couplé. Soient $\delta t > 0$ le pas de temps, et $t^n = n\delta t$, $n \in \mathbb{N}$. Notons $\mathbf{u}^n$ et p^n la solution approchée au temps t^n. Le terme intégral $\displaystyle\int_0^{t^n} \int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n}$ se discrétise comme suit

$$\int_0^{t^n} \int_{\Gamma_0} \mathbf{u}\cdot\mathbf{n} \approx \sum_{j=0}^{n} \left(\delta t \int_{\Gamma_0} \mathbf{u}^j \cdot \mathbf{n}\right). \tag{3.13}$$

Le système discrétisé en temps s'écrit donc

$$\left\{\begin{array}{rcll} \rho\dfrac{\mathbf{u}^n}{\delta t} - \eta\triangle\mathbf{u}^n + \nabla p^n &=& \rho\dfrac{\mathbf{u}^{n-1}}{\delta t}, & \text{sur } \Omega, \\ \nabla\cdot\mathbf{u}^n &=& 0, & \text{sur } \Omega, \\ \mathbf{u}^n &=& 0, & \text{sur } \Gamma_\ell, \\ \eta\nabla\mathbf{u}^n\cdot\mathbf{n} - p^n\mathbf{n} &=& -\left[P_0 + \dfrac{f^n_{ext}}{S} - \dfrac{kx_0}{S} + \dfrac{k\delta t}{S^2}\displaystyle\sum_{j=0}^{n-1}\left(\int_{\Gamma_0}\mathbf{u}^j\cdot\mathbf{n}\right) - \dfrac{m}{S^2\delta t}\left(\int_{\Gamma_0}\mathbf{u}^{n-1}\cdot\mathbf{n}\right)\right]\mathbf{n} & \\ & & -\left(\dfrac{m}{S^2\delta t} - \dfrac{k\delta t}{S^2} + \dfrac{\mu}{S^2}\right)\left(\displaystyle\int_{\Gamma_0}\mathbf{u}^n\cdot\mathbf{n}\right)\mathbf{n}, & \text{sur } \Gamma_0, \\ \eta\nabla\mathbf{u}^n\cdot\mathbf{n} - p^n\mathbf{n} &=& -R_i\left(\displaystyle\int_{\Gamma_i}\mathbf{u}^n\cdot\mathbf{n}\right)\mathbf{n}, & \text{sur } \Gamma_i,\ i = 1,\dots,N. \end{array}\right. \tag{3.14}$$

Dans la suite, nous utiliserons les notations suivantes

$$F^n = P_0 + \frac{f^n_{ext}}{S} - \frac{kx_0}{S} + \frac{k\delta t}{S^2}\sum_{j=0}^{n-1}\left(\int_{\Gamma_0}\mathbf{u}^j\cdot\mathbf{n}\right) - \frac{m}{S^2\delta t}\left(\int_{\Gamma_0}\mathbf{u}^{n-1}\cdot\mathbf{n}\right) \tag{3.15}$$

et

$$R_0^{\delta t} = \frac{m}{S^2\delta t} - \frac{k\delta t}{S^2} + \frac{\mu}{S^2}. \tag{3.16}$$

Notons que F^n est connue dès que les vitesses au pas de temps précédent ont été calculées. Le problème (3.14) peut être réécrit ainsi

$$\left\{\begin{array}{rcll} \rho\dfrac{\mathbf{u}^n}{\delta t} - \eta\triangle\mathbf{u}^n + \nabla p^n &=& \rho\dfrac{\mathbf{u}^{n-1}}{\delta t}, & \text{sur } \Omega, \\ \nabla\cdot\mathbf{u}^n &=& 0\,, & \text{sur } \Omega, \\ \mathbf{u}^n &=& 0, & \text{sur } \Gamma_\ell, \\ \eta\nabla\mathbf{u}^n\cdot\mathbf{n} - p^n\mathbf{n} &=& -F^n\mathbf{n} - R_0^{\delta t}\left(\displaystyle\int_{\Gamma_0}\mathbf{u}^n\cdot\mathbf{n}\right)\mathbf{n}\,, & \text{sur } \Gamma_0, \\ \eta\nabla\mathbf{u}^n\cdot\mathbf{n} - p^n\mathbf{n} &=& -R_i\left(\displaystyle\int_{\Gamma_i}\mathbf{u}^n\cdot\mathbf{n}\right)\mathbf{n}, & \text{sur } \Gamma_i,\ i = 1,\dots,N. \end{array}\right. \tag{3.17}$$

En conséquence, après discrétisation en temps, nous obtenons un système de Stokes généralisé avec des conditions aux bords naturellement dissipatives. Ces conditions aux bords inhabituelles provenant des résistances et de la discrétisation en temps du système masse–ressort modifient la forme bilinéaire standard associée à un problème de Stokes avec conditions aux bords mixtes de Dirichlet–Neumann. En particulier, si on considère une discrétisation en éléments finis, elles coupleraient tous les degrés de liberté de chaque sortie et modifieraient la forme de matrice éléments finis associée aux degrés de liberté de la vitesse du fluide. Par suite, ces conditions aux bords ne peuvent pas être aisément et directement implémentées dans un solveur éléments finis quelqu'il soit, par exemple `FreeFEM++`, sans rentrer profondément dans le code.

Afin de passer outre cette difficulté, nous utilisons la méthode suivante. L'objectif est de réduire ce problème muni de conditions aux bords non–standard à plusieurs autres problèmes munis cette fois de conditions de Neumann et Dirichlet classiques. L'idée est de pré–calculer un jeu de solutions avec conditions de Neumann sur chaque Γ_i, puis de définir la solution comme combinaison linéaire de ces solutions et d'un terme correctif. Ce terme correctif est également solution d'un problème de Stokes avec conditions aux bords standard et les coefficients de la combinaison linéaire sont calculés de façon à ce que les conditions aux bords dissipatives soient satisfaites.

Plus précisément, grâce à la linéarité du système à chaque pas de temps, la solution $(\mathbf{u}^n, p^n)$ de (3.17) peut se décomposer selon

$$\left\{\begin{array}{rcll} \mathbf{u}^n &=& \tilde{\mathbf{u}}^n + \displaystyle\sum_{i=0}^{N}\alpha_i^n\mathbf{u}_i & \text{et} \\ p^n &=& \tilde{p}^n + \displaystyle\sum_{i=0}^{N}\alpha_i^n p_i, & \end{array}\right. \tag{3.18}$$

où $(\mathbf{u}_i)_{i=0}^N$ et $(p_i)_{i=0}^N$ ne dépendent pas du temps, à l'inverse de $\tilde{\mathbf{u}}^{n+1}$ et $\tilde{p}^{n+1}$ qui sont les termes correctifs calculés à chaque pas de temps :

- pour tout $i = 0, \dots N$, la solution $(\mathbf{u}_i, p_i)$ est calculée durant **l'étape de pré–traitement** (3.19)

$$\left\{\begin{array}{rcll} \rho\dfrac{\mathbf{u}_i}{\delta t} - \eta\triangle\mathbf{u}_i + \nabla p_i &=& 0\,, & \text{sur } \Omega\,, \\ \nabla\cdot\mathbf{u}_i &=& 0\,, & \text{sur } \Omega\,, \\ \mathbf{u}_i &=& 0, & \text{sur } \Gamma_\ell\,, \\ \eta\nabla\mathbf{u}_i\cdot\mathbf{n} - p_i\mathbf{n} &=& 0\,, & \text{sur } \Gamma_j\,,\ \text{pour } j\neq i\,, \\ \eta\nabla\mathbf{u}_i\cdot\mathbf{n} - p_i\mathbf{n} &=& -\mathbf{n}\,, & \text{sur } \Gamma_i. \end{array}\right. \tag{3.19}$$

- Puis, à chaque pas de temps, $\tilde{\mathbf{u}}^n$ prend en compte le terme non–stationnaire

$$\left\{\begin{array}{rcll} \rho\dfrac{\tilde{\mathbf{u}}^n}{\delta t} - \eta\triangle\tilde{\mathbf{u}}^n + \nabla\tilde{p}^n &=& \rho\dfrac{\mathbf{u}^{n-1}}{\delta t}\,, & \text{sur } \Omega\,, \\ \nabla\cdot\tilde{\mathbf{u}}^n &=& 0\,, & \text{sur } \Omega\,, \\ \tilde{\mathbf{u}}^n &=& 0, & \text{sur } \Gamma_\ell\,, \\ \eta\nabla\tilde{\mathbf{u}}^n\cdot\mathbf{n} - \tilde{p}^n\mathbf{n} &=& 0\,, & \text{sur } \Gamma_i \quad i = 0, \dots, N. \end{array}\right. \tag{3.20}$$

- Enfin les coefficients α_i^n sont choisis tels que les conditions sur les sorties Γ_i soient vérifiées. Cela implique que les α_i^n, $i = 0, \dots, N$ soient solution d'un **système linéaire**.

Système linéaire vérifié par les $(\alpha_i^n)_{0\le i\le N}$:

– Sur Γ_0, nous avons $(\eta\nabla\mathbf{u}^n - p^n)\,\mathbf{n} = -F^n\,\mathbf{n} - R_0^{\delta t}\left(\displaystyle\int_{\Gamma_0}\mathbf{u}^n\cdot\mathbf{n}\right)\mathbf{n}$. À partir de la relation (3.18), nous obtenons

$$(\eta\nabla\tilde{\mathbf{u}}^n - \tilde{p}^n)\,\mathbf{n} + \sum_{i=0}^N \alpha_i^n\,(\eta\nabla\mathbf{u}_i - p_i)\,\mathbf{n} = -F^n\,\mathbf{n} - R_0^{\delta t}\left(\int_{\Gamma_0}\tilde{\mathbf{u}}^n\cdot\mathbf{n}\right)\mathbf{n} - R_0^{\delta t}\sum_{i=0}^N \alpha_i^n\left(\int_{\Gamma_0}\mathbf{u}_i\cdot\mathbf{n}\right)\mathbf{n}.$$

Grâce aux conditions aux bords des problèmes (3.19) et (3.20), nous trouvons

$$\alpha_0^n = F^n + R_0^{\delta t}\left(\int_{\Gamma_0}\tilde{\mathbf{u}}^n\cdot\mathbf{n}\right) + R_0^{\delta t}\sum_{i=0}^N \alpha_i^n\left(\int_{\Gamma_0}\mathbf{u}_i\cdot\mathbf{n}\right). \tag{3.21}$$

– Sur Γ_j, $j\neq 0$, de la même façon, nous obtenons

$$\alpha_j^n = R_j\left(\int_{\Gamma_j}\tilde{\mathbf{u}}^n\cdot\mathbf{n}\right) + R_j\sum_{i=0}^N \alpha_i^n\left(\int_{\Gamma_j}\mathbf{u}_i\cdot\mathbf{n}\right). \tag{3.22}$$

Soit $\mathbf{b}^n = (b_j^n)_{0\le j\le N} \in \mathbb{R}^{N+1}$ le vecteur défini par

$$\left\{\begin{array}{rcl} b_0^n &=& F^n + R_0^{\delta t}\left(\displaystyle\int_{\Gamma_0}\tilde{\mathbf{u}}^n\cdot\mathbf{n}\right),\ \text{et} \\ b_j^n &=& R_j\left(\displaystyle\int_{\Gamma_j}\tilde{\mathbf{u}}^n\cdot\mathbf{n}\right),\ \text{pour } j = 1, \dots, N, \end{array}\right. \tag{3.23}$$

et $A = (a_{ij})_{0\le i,j\le N}$ la matrice

$$a_{ij} = \tilde{R}_i\left(\int_{\Gamma_i}\mathbf{u}_j\cdot\mathbf{n}\right),\ \text{ pour } i, j = 0, \dots, N, \tag{3.24}$$

où $\tilde{R}_0 = R_0^{\delta t}$ et $\tilde{R}_i = R_i$ si $i \neq 0$.

Par suite, à chaque itération en temps, $(\alpha_i^n)_{0\leq i\leq N}$ est solution du système linéaire suivant.

$$(I - A) \cdot \mathbf{x} = \mathbf{b}^n. \tag{3.25}$$

Proposition 1 *La matrice* $(I - A)$ *est inversible.*

Preuve.

1. Prenons $v \in V$ (voir (3.4)), une fonction test. La formulation variationnelle pour les pré–calculs s'écrit :

$$\frac{\rho}{\delta t}\int_\Omega \mathbf{u}_i \cdot \mathbf{v} + \eta \int_\Omega \nabla \mathbf{u}_i \cdot \nabla \mathbf{v} = -\int_{\Gamma_i} \mathbf{v} \cdot \mathbf{n}, \qquad \forall i = 1, \ldots, N. \tag{3.26}$$

   En prenant successivement, $\mathbf{v} = \mathbf{u}_j$ dans (3.26) écrite pour $\mathbf{u}_i$ et $\mathbf{v} = \mathbf{u}_i$ dans celle écrite pour $\mathbf{u}_j$, nous obtenons facilement que

$$\int_{\Gamma_i} \mathbf{u}_j \cdot \mathbf{n} = \int_{\Gamma_j} \mathbf{u}_i \cdot \mathbf{n}, \qquad \forall (i, j) \in [\![1, N]\!]^2.$$

   De plus la forme bilinéaire

$$\begin{array}{rcl} V \times V & \to & \mathbb{R} \\ (\mathbf{u}, \mathbf{v}) & \mapsto & \dfrac{\rho}{\delta t}\displaystyle\int_\Omega \mathbf{u} \cdot \mathbf{v} + \eta \int_\Omega \nabla \mathbf{u} \cdot \nabla \mathbf{v} \end{array}$$

   est clairement positive et définie. Cela montre que la matrice

$$(-Q)_{i,j} = -\int_{\Gamma_i} \mathbf{u}_j \cdot \mathbf{n}$$

   est symétrique définie positive.

2. Soit $R = \mathrm{diag}[R_0, \ldots, R_n]$. Les R_i sont strictement positifs ; notons $\sqrt{R}$ sa racine carrée. Nous avons l'identité suivante :

$$(I - RQ) = \sqrt{R}\left(I - \sqrt{R}\, Q\, \sqrt{R}\right) \sqrt{R}^{-1}.$$

   $(I - RQ)$ est donc semblable à une matrice symétrique définie positive, donc est inversible. Nous terminons en remarquant que $RQ = A$.

□

Remarque 9 *La méthode utilisée ici est fondée sur la linéarité du problème. Néanmoins, elle nous permet de traiter ces conditions aux bords non–standard en résolvant uniquement des problèmes aux conditions aux bords classiques.*

3.3.2 Extension au problème de Navier–Stokes

Nous nous intéressons maintenant au problème initial (3.3), *i.e* nous ajoutons un terme de convection $(\mathbf{u} \cdot \nabla)\mathbf{u}$ dans l'équation du moment pour le fluide. Bien que la méthode précédemment décrite ne soit valable que pour des problèmes linéaires, elle peut être adaptée. Nous utilisons la même étape de pré–traitement, calculant les solutions $(\mathbf{u}_i)_{i=0}^N$ des problèmes de Stokes (3.12). Afin de pouvoir appliquer les mêmes techniques que celles décrites pour le problème de Stokes, nous traitons la non–linéarité soit explicitement soit par la méthode des caractéristiques. La première nécessite un petit pas de temps, c'est pourquoi nous choisissons la seconde. La discrétisation de la première ligne de (3.17) devient [Pir88]

$$\rho \frac{\tilde{\mathbf{u}}^n}{\delta t} - \eta \triangle \tilde{\mathbf{u}}^n + \nabla \tilde{p}^n \quad = \quad \frac{\rho}{\delta t}\, \mathbf{u}^{n-1} \circ X^{n-1}, \quad \text{sur } \Omega, \tag{3.27}$$

où $X^{n-1}(\mathbf{x})$ est une approximation du pied de la caractéristique au temps $(n-1)\,\delta t$ qui passe par $\mathbf{x}$ au temps $n\,\delta t$ ($X^{n-1} \approx \mathbf{x} - \mathbf{u}^{n-1}(\mathbf{x})\delta t$) et $\mathbf{x}$ est le point courant dans $\mathbb{R}^d$. Sous une condition de type CFL, ce schéma est inconditionnellement stable et convergent (voir [Pir88], page 147). En pratique, nous utiliserons la fonction `convect` déjà implémentée en `Freefem++` [HLHOP].

Concernant le terme correctif $\tilde{\mathbf{u}}^n$, la résolution peut être réalisée par différentes méthodes, par exemple par des méthodes fondées sur une formulation variationnelle mixte obtenue en multipliant (3.27) par $(\mathbf{v}, q) \in H^1(\Omega)^d \times L^2(\Omega)$, où $\mathbf{v}$ s'annule sur Γ_ℓ. Grâce à l'incompressibilité et aux conditions aux bords, nous obtenons

$$\left\{\begin{array}{rcl} \dfrac{\rho}{\delta t}\displaystyle\int_\Omega \tilde{\mathbf{u}}^n \cdot \mathbf{v} - \eta \int_\Omega \nabla\tilde{\mathbf{u}}^n : \nabla\mathbf{v} - \int_\Omega \nabla\cdot\mathbf{v}\tilde{p}^n &=& \dfrac{\rho}{\delta t}\displaystyle\int_\Omega \mathbf{u}^{n-1}\circ X^{n-1}\cdot\mathbf{v}, \\ -\displaystyle\int_\Omega q\nabla\cdot\tilde{\mathbf{u}}^n &=& 0\,, \qquad \forall q \in L^2(\Omega). \end{array}\right. \tag{3.28}$$

Comme nous venons de le mentionner, le second membre de (3.28) est calculé à l'aide de la fonction `convect` déjà implémentée dans `Freefem++`[HLHOP].

Afin d'accélérer les calculs nous avons ensuite choisi de calculer le terme correctif à l'aide d'une méthode de projection. La méthode de projection de CHORIN [Cho68],[Cho69] (voir également TEMAM [Tem77] et QUARTAPELLE [Qua93]) est fréquemment utilisée pour la résolution des équations de Navier–Stokes en variables primitives. Cette méthode de splitting est fondée sur une discrétisation en temps des équations décrivant le comportement de fluides visqueux incompressibles, dans laquelle la viscosité et l'incompressibilité du fluide sont traitées en deux étapes distinctes. Ceci permet de découpler le calcul de la pression du fluide de celui de la vitesse du fluide. Dans un premier temps, nous choisissons de résoudre le problème par une méthode mixte. Nous détaillons plus en avant les méthodes de projections que nous avons implémentées sur ce problème à la Section 3.3.3.

3.3.3 Méthodes de projection

Une des principales difficultés liées à la résolution des équations de Navier–Stokes est le couplage entre l'inconnue de vitesse $\tilde{\mathbf{u}}^n$ et de pression p par la contrainte d'incompressibilité. Les méthodes de projection introduites par CHORIN et TEMAM [Cho68, Cho69] pour approcher numériquement les solutions sont connues pour leur simplicité et leur efficacité. Ces méthodes proposent le découplage des approximations de la vitesse et de la pression à chaque pas de temps, évitant ainsi les difficultés inhérentes à la résolution du problème de Stokes. Dans un premier temps nous résolvons un problème de convection–diffusion tel que la vitesse résultante ne soit pas à divergence nulle, puis dans un deuxième temps celle–ci est projetée sur un espace de fonctions à divergence nulle afin de satisfaire la contrainte d'incompressibilité.

Depuis leur introduction par CHORIN et TEMAM, les méthodes de projection ont fait l'objet de nombreux travaux et il existe de nombreuses variantes. Pour rendre notre code plus efficace, nous en avons implémenté une version très simple.

Méthode de projection non–incrémentale

Au lieu de résoudre le problème de Navier–Stokes pour le terme correctif $\tilde{\mathbf{u}}^n$, nous appliquons une méthode de projection. L'idée, commune à tous les schémas de projection, est de découpler les effets visqueux du fluide de l'incompressibilité de l'écoulement ; ils sont ainsi fondés sur une stratégie de "prédiction–correction". Introduisons donc deux suites $(\bar{\mathbf{u}}^n)_n$ et $(\mathbf{u}_p^n)_n$ pour la vitesse et (p_p^n) pour la pression telles que à chaque pas de temps :

- ⋆ $\mathbf{u}_p^n$ soit une "prédiction" du champ de vitesse, vérifiant les conditions limites du problème (3.20) mais ne vérifiant pas nécessairement la contrainte de divergence nulle ;
- ⋆ $\bar{\mathbf{u}}^n$ une version "corrigée" de $\mathbf{u}_p^n$, à divergence nulle cette fois, mais ne vérifiant pas forcément les conditions aux limites de (3.20) ;
- ⋆ p_p^n permet de construire la "correction" pour obtenir $\bar{\mathbf{u}}^n$ et correspond à l'approximation de la pression. Soulignons qu'obtenir $\bar{\mathbf{u}}^n$ revient à calculer p_p^n.

Nous détaillons ci–dessous l'algorithme qu'on effectue à chaque pas de temps :

Étape 1 Nous résolvons :

$$\begin{array}{rcl} \frac{\rho}{\Delta t}(\mathbf{u}_p^n - \mathbf{u}^{n-1} \circ X^{n-1}) - \eta\Delta\mathbf{u}_p^n &=& 0, \quad \text{dans} \quad \Omega, \\ \mathbf{u}_p^n &=& 0, \quad \text{sur} \quad \Gamma_\ell, \\ \eta\nabla\mathbf{u}_p^n \cdot \mathbf{n} &=& 0, \quad \text{sur} \quad \Gamma_i,\ i = 0, \ldots, N, \end{array} \tag{3.29}$$

de fonction inconnue $\mathbf{u}_p^n$. C'est un problème elliptique. Notons que le terme de pression n'apparaît plus dans cette étape.

Étape 2 Nous calculons le terme correctif p_p^n solution du problème elliptique suivant :

$$\begin{array}{rcll} -\Delta p_p^n &=& -\frac{\rho}{\delta t}\nabla \cdot \mathbf{u}_p^n, & \text{dans} \quad \Omega, \\ \nabla p_p^n \cdot \mathbf{n} &=& 0, & \text{sur} \quad \Gamma_\ell, \\ p_p^n &=& 0, & \text{sur} \quad \Gamma_i,\ i = 0, \ldots, N. \end{array} \tag{3.30}$$

Étape 3 ⋆ Nous corrigeons :

$$\bar{\mathbf{u}}^n = \mathbf{u}_p^n - \frac{\delta t}{\rho}\nabla p_p^n$$

qui est maintenant à divergence nulle.

⋆ Les α_i sont calculés de façon à ce que

$$\mathbf{u}^n = \bar{\mathbf{u}}^n + \sum_{i=0}^{N} \alpha_i \mathbf{u}_i$$

vérifie les conditions aux bords, où les $\mathbf{u}_i$ sont les solutions en vitesse des pré–calculs (3.19). C'est–à–dire que les α_i sont les solutions du système linéaire (3.25), où le vect $\mathbf{b}^n$ (3.23) est calculé en remplaçant $\tilde{u}^n$ par $\bar{u}^n$ dans la formule (3.23). Nous calculons également la pression :

$$p^n = p_p^n + \sum_{i=0}^{N} \alpha_i p_i,$$

avec p_i également les solutions en pression des pré–calculs

Remarque 10 *Nous avons choisi de résoudre ici l'étape 3 de projection en considérant une formulation de Poisson sur la pression. Cette étape peut néanmoins se formuler à l'aide d'un système de Darcy en vitesse–pression. Cependant la formulation de Poisson permet d'accéler les calculs et de s'affranchir des difficultés liées aux problèmes mixtes (en particulier la condition inf–sup [BF91]).*

Méthode de projection incrémentale

La version non–incrémentale du schéma de Chorin–Teman est connue pour ne pas être d'ordre 1 en temps. Afin d'avoir un schéma d'ordre plus élevé nous avons également implémenté sa version incrémentale. Dans la méthode précédemment décrite le terme de pression est recalculé à chaque étape. Celui–ci n'apparaît pas dans le calcul de la première étape (3.29). Dans cette nouvelle méthode, nous adaptons la méthode précédente en utilisant la pression calculée au pas de temps précédent comme estimation de la nouvelle pression. Nous avons ainsi deux systèmes elliptiques à résoudre.

Nous détaillons ci–dessous l'algorithme qu'on effectue à chaque pas de temps :

Étape 1 Nous résolvons :

$$\begin{array}{rcl} \frac{\rho}{\Delta t}(\mathbf{u}_p^n - \mathbf{u}^{n-1} \circ X^{n-1}) - \eta\Delta\mathbf{u}_p^n + \nabla p^{n-1} &=& 0, \quad \text{dans} \quad \Omega, \\ \mathbf{u}_p^n &=& 0, \quad \text{sur} \quad \Gamma_\ell, \\ \eta\nabla\mathbf{u}_p^n \cdot \mathbf{n} &=& 0, \quad \text{sur} \quad \Gamma_i,\ i = 0, \ldots, N, \end{array} \tag{3.31}$$

de fonction inconnue $\mathbf{u}_p^n$. C'est à nouveau un problème elliptique. Notons l'ajout du terme de pression ∇p^{n-1} par rapport à (3.29) donné par le calcul au pas de temps précédent.

Étape 2 Nous calculons le terme correctif p_p^n solution du problème elliptique suivant :

$$\begin{array}{rcll} -\Delta p_p^n & = & -\dfrac{\rho}{\delta t}\nabla\cdot\mathbf{u}_p^n, & \text{dans} \quad \Omega, \\ \nabla p_p^n\cdot\mathbf{n} & = & 0, & \text{sur} \quad \Gamma_\ell, \\ p_p^n & = & 0, & \text{sur} \quad \Gamma_i,\ i=0,\dots,N. \end{array} \tag{3.32}$$

Cette étape est identique à l'étape 2 de la méthode non–incrémentale.

Étape 3 ⋆ Nous corrigeons la pression,

$$\bar{p}^n = p^{n-1} + p_p^n$$

et la vitesse

$$\bar{\mathbf{u}}^n = \mathbf{u}_p^n - \frac{\delta t}{\rho}\nabla p_p^n$$

qui est maintenant à divergence nulle.

⋆ Comme pour la méthode de projection non–incrémentale, les α_i sont calculés de façon à ce que

$$\mathbf{u}^n = \bar{\mathbf{u}}^n + \sum_{i=0}^{N}\alpha_i\mathbf{u}_i$$

vérifie les conditions aux bords. Calculons également la pression :

$$p^n = \bar{p}^n + \sum_{i=0}^{N}\alpha_i p_i.$$

3.3.4 Pertinence du choix d'une méthode de projection

Afin de comparer les temps de calculs pour les différentes méthodes décrites précédemment, nous lançons des calculs sur un domaine 2D décrivant un arbre composé de 3 générations pour un régime de respiration normal. Nous comparons les solutions approchées obtenues au temps $t = 5\,\text{s}$. Le pas de temps est fixe pour les 3 calculs, $\delta t = 0.01\,\text{s}$. Afin de comparer les deux m/'ethodes, nous avons utilisé dans les deux cas pour la discrétisation en espace du système de Navier–Stokes des mini–éléments P1/P2. La Table 3.1 représente les temps de calculs CPU pour chacune des méthodes ainsi que l'erreur en norme L^2 en pression par rapport à une solution de référence calculée sur un mailage fin pour un pas de temps petit.

Méthode	Temps CPU	$\frac{\|p-p_{ref}\|_2}{\|p_{ref}\|_2}$
Mixte	1962.61	0.033
Projection incrémentale	1630.46	0.05
Projection non–incrémentale	1626.10	0.02

TAB. 3.1: Comparaison des temps de calcul CPU pour la méthode mixte et les méthodes de projection

3.4 Simulations numériques

Les méthodes décrites plus tôt ont été implémentées en `FreeFEM++` (voir [HLHOP]). Nous avons utilisé les mini–éléments P1/P2 pour la discrétisation en espace du système de Navier–Stokes, voir par exemple [ABF84]. Nous avons réalisé des simulations numériques dans un arbre bronchique 2D idéalisé (jusqu'à la 3e génération, voir par exemple Figure 3.9). Notons tout de même que les quantités géométriques de cet arbre 2D sont choisies telles que les résistances de cet arbre 2D soient égales à celles d'un arbre réaliste 3D (voir par exemple [SA07]). Afin de conserver ces résistances, nous modifions également la viscosité du fluide (voir ci–dessous).

Taille du maillage ou geométrie : nous utilisons une triangulation composée de

- $N_{vertices} = 9996$ noeuds,
- $N_{triang} = 17340$ triangles et
- $N_{edges} = 2650$ arêtes.
- La géométrie de l'arbre est contenue dans une boîte $[-0,22939,\ 0,24739] \times [-0,35569, 0]$.

Données physiologiques (paramètres pour le fluide et le ressort) : Afin de reproduire des courbes expérimentales, nous avons utilisé des données physiologiques, choisies proches de celles que l'on trouve dans [BT06], [MMSS08]

- $m = 0.3\,\mathrm{kg}$, la masse totale du poumon ;
- $S = 0.011\,\mathrm{m}^2$, surface de la face mobile de la boîte (surface du diaphragme) ;
- $E = 3.32 \times 10^5\,\mathrm{N \cdot m^{-5}}$, l'élastance du poumon ;
- $k_{00} = E \cdot S^2 = 40.172\,\mathrm{N \cdot m^{-1}}$, la constante de raideur du ressort ;
- $R_{out} = 1.33 \cdot 10^5\,\mathrm{Pa \cdot s \cdot m^{-3}}$,la résistance à chaque sortie $\Gamma_i,\ i \geq 0$;
- $R_{in} = 1.12 \times 10^5\,\mathrm{Pa \cdot s \cdot m^{-3}}$, la résistance à l'entrée (trachée) ;

Concernant la viscosité du fluide η et la densité du fluide ρ, nous ne choisissons pas les valeurs physiologiques mais des valeurs telle que la résistance globale de notre arbre 2D soit égale à la résistance globale de l'arbre réel 3D, et telles que le nombre de Reynolds 2D soit égal au nombre de Reynolds 3D (voir [SA07]). En conséquence,

- $\eta = 0.004\,\mathrm{Pa \cdot s \cdot m}$,
- $\rho = 50\,\mathrm{kg \cdot m^{-3}}$.

Le fait de conserver les résistances, permet de conserver les débits, et conserver le nombre de Reynolds permet de conserver tout phénomène de circulation. Toutefois, on ne peut pas conserver toutes les caractéristiques du fluide ainsi. En particulier ici, on ne conserve pas la norme de la vitesse. Cependant, le comportement qualitatif de la vitesse est conservé.

3.4.1 Premier modèle

Nous réalisons les calculs dans un cas où il n'y a pas de force extérieure. Dans un premier temps, la résistance à l'entrée est égale à R_{in} et les résistances en sortie R_i sont toutes prises égales à R_{out}. De plus, la constante de raideur du ressort k est fixée constante égale à k_0. Initialement, nous supposons que le ressort est étiré de sorte que $x(0) = 0.1\,\mathrm{m}$. Il est lâché sans vitesse initiale. La Figure 3.1 représente les variations du volume $V = Sx$ et le flux d'air à travers la frontière Γ_0 en fonction du temps.

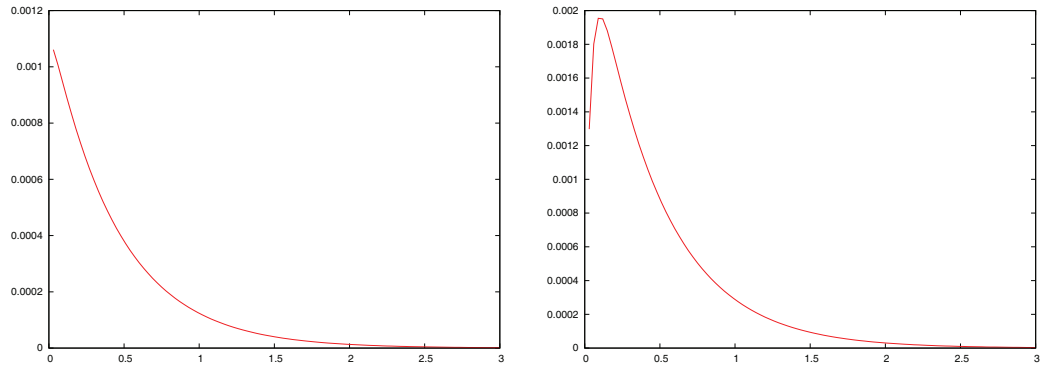
Afin de visualiser l'influence du coefficient de raideur du ressort sur le système, nous traçons la variation de volume $V = Sx$ et le flux d'air à travers Γ_0 en fonction du temps pour différentes valeurs de k_0, $35 < k_0 < 75\,\mathrm{N \cdot m^{-1}}$ (voir Figure 3.2). Nous réalisons également des simulations en faisant varier cette fois la valeur des résistances R_i à chaque sortie Γ_i, avec $k = k_0$, constant.

Commentaires :

- Lorsque la constante de raideur k est grande, le flux d'air maximal est plus grand et le ressort retourne à sa position d'équilibre plus rapidement (voir Figure 3.2).
- Remarquons que la modification des résistances sur uniquement deux sorties (la 7e et la 8e) ne semble pas modifier significativement la relaxation du ressort (voir Figure 3.3).

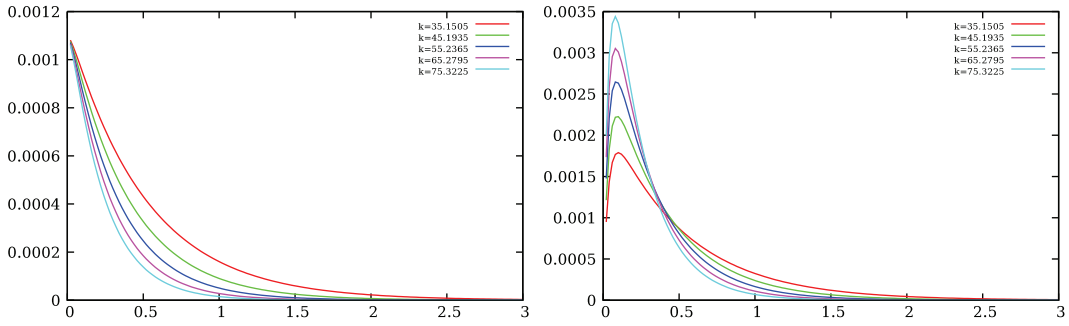
3.4.2 Respiration forcée

Nous présentons ici les résultats numériques obtenus dans le cas de manoeuvres forcées. Ces manoeuvres sont habituellement réalisées par les patients afin de déterminer les pathologies dont ils souffrent. Dans ce cas, la force f_{ext} s'exerçant sur le ressort est donnée à la Figure 3.4. Notons que la première partie de la manoeuvre, pour $0 < t < 8\,\mathrm{s}$ correspond à la **respiration au repos**. Afin de simuler la respiration d'un patient réalisant une



(a) Volume (b) Flux

FIG. 3.1: Variation du volume $V = Sx$ (en m^3) et flux d'air (en $m^3 \cdot s^{-1}$) à travers la frontière Γ_0 en fonction du temps (en s), $k \equiv k_0$.



(a) Volume (b) Flux

FIG. 3.2: Variation du volume $V = Sx$ (en m^3) et flux d'air (en $m^3 \cdot s^{-1}$) à travers la frontière Γ_0 en fonction du temps (en s) pour différentes valeurs de k.

manoeuvre forcée et dans le but de reproduire des données physiologiques, nous choisissons de prendre pour nos simulations numériques, une constante de raideur et des résistances dépendant du volume d'air dans le poumon, *i.e.* dépendant de x, comme décrit à la Section 2.4.1, page 84. Les valeurs numériques choisies pour les paramètres sont données à l'Annexe 5.4, page 136.

La force extérieure f_{ext} est choisie constante par morceaux, avec

$$f_{min} \leq f_{ext} \leq f_{max}\ ,\ f_{min} = -11\,\mathrm{N} \text{ and } f_{max} = 13\,\mathrm{N} \quad \text{(Voir Figure 3.4)},$$

et afin de prendre en compte l'influence du muscle lisse, nous choisissons

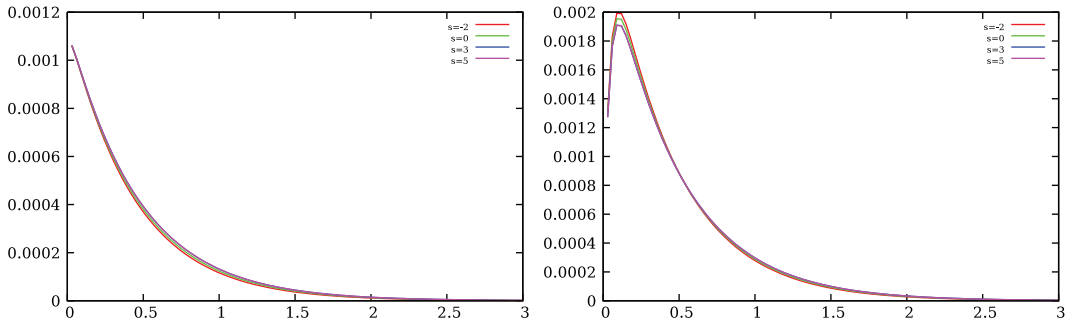
$$R_i(x) = \frac{R_i}{1 + \theta S x / V_B^0}, \quad i = 1, \ldots, N, \tag{3.33}$$

où

- R_i est la résistance des voies respiratoires du i^e sous–arbre, qui mesure les forces résistives du sous–arbre bronchique correspondant.
- Le paramètre θ décrit l'activité du muscle lisse, voir Section 2.4.1, page 85

Dans ce qui suit, V_A^0 (resp. V_B^0) désigne le volume d'air dans les alvéoles (resp. dans les bronches) au repos.(Voir à nouveau [SA07] pour plus de détails sur ce modèle.)

Le paramètre de répartition est par défaut égal à $\theta = 0.4$ dans un poumon humain. Quand il est plus faible, cela indique une activité du muscle lisse importante : en effet, un θ faible induit de plus faibles variations du volume bronchique, c'est–à–dire une activité importante du muscle lisse.



(a) Volume (b) Flux

FIG. 3.3: Variation du volume $V = Sx$ (en m^3) et flux d'air (en $m^3 \cdot s^{-1}$) à travers la frontière Γ_0 en fonction du temps (en s) pour différentes valeurs de R, $R_i = R_0$, pour $i = 1, \ldots, 6$ et $R_7 = R_8 = R_0 \times 10^s$, $s = -2, 0, 3, 5$, $k = k_0$.

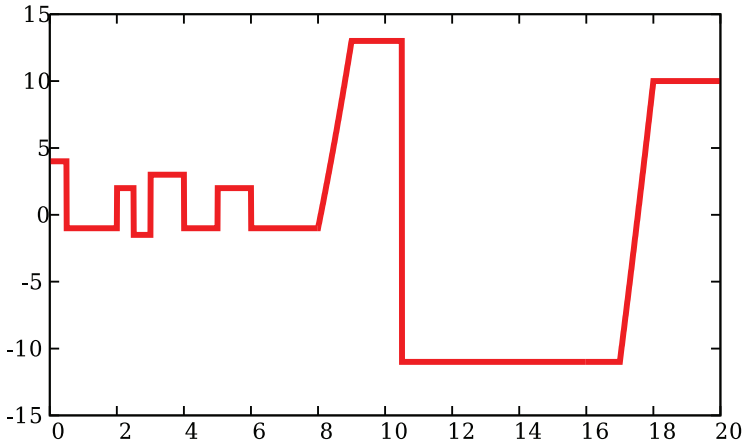


FIG. 3.4: Force extérieure f_{ext} (en N) en fonction du temps (en s).

Nous choisissons également une force de retour élastique non–linéaire, suivant le modèle présenté Section 2.4.1, page 84.

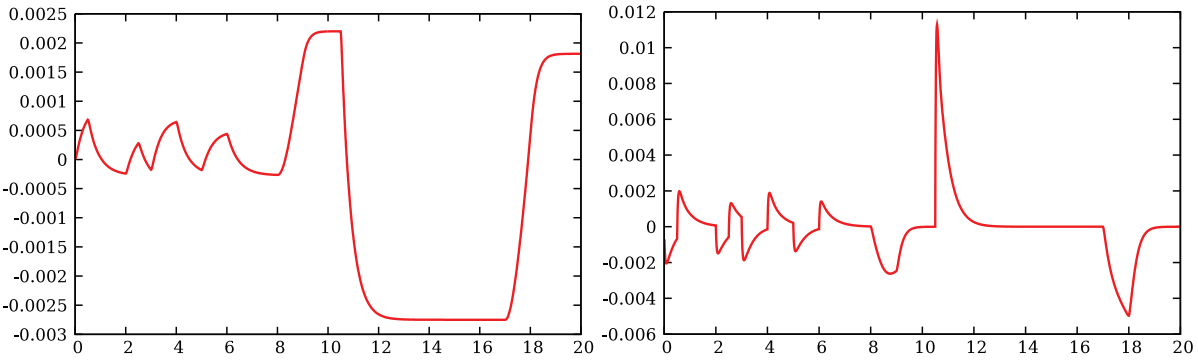
3.4.3 Un cas de référence : données non pathologiques

Dans le cas où la force extérieure f_{ext} décrite à la Figure 3.4 est exercée, nous calculons les variations du volume d'air pulmonaire et le flux d'air à la bouche (voir Figure 3.10) pour $\theta = 0.011$, $R_i = R_{out}$ et $k_0 = k_{00}$. On trace également la norme de la vitesse de l'air et la pression dans l'arbre bronchique (voir Figure 3.9) au temps $t = 20$s. Le portrait de phase (voir Figure 3.11) est typiquement la courbe obtenue par spirométrie que les pneumologues ont l'habitude de lire. C'est ce genre de courbes qui leur apporte les informations nécessaires au diagnostic. (Voir Section 2.2.3, page 76.)

Qualitativement, nous retrouvons les courbes obtenues par spirométrie. Pour s'en convaincre comparer la Figure 3.11 au résultat d'examen médical Figure 2.6 ou bien au résultat d'examen non pathologique donné dans [MMSS08]. Insistons sur le fait que nous retrouvons les bons ordres de grandeur pour les volumes et débits.

3.4.4 Étude de la sensitivité

Afin de mieux comprendre les pathologies du poumon, nous voulons observer l'influence de la raideur du ressort, du paramètre θ et des résistances, en particulier sur le portrait de phase. Tout d'abord nous choisissons de prendre différentes valeurs pour le coefficient de raideur k (voir Figure 3.7), avec $\theta = 0.011$ et $R_i = R_{out}$.



(a) Volume (b) Flux

FIG. 3.5: Variations de volume (en m^3) et flux d'air (en $m^3 \cdot s^{-1}$) à travers le bord Γ_0 en fonction du temps en régime forcé avec les données physiologiques correspondant à un patient normal.

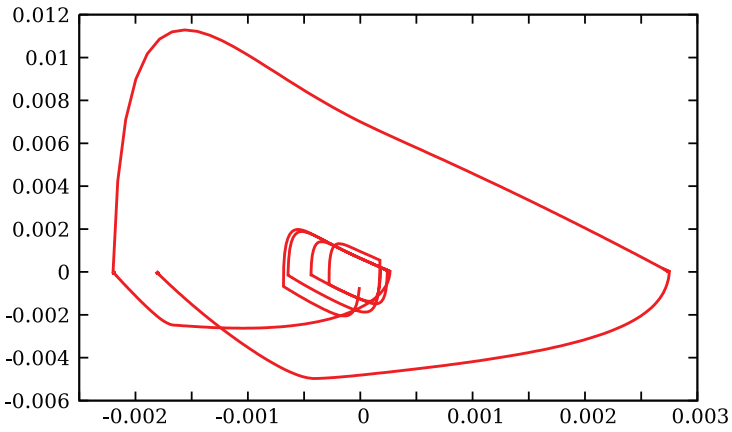


FIG. 3.6: Portrait de phase flux (en $m^3 \cdot s^{-1}$) – volume (en m^3).

Le coefficient k correspond physiologiquement dans notre modèle à la raideur bronchique. Une diminution de ce paramètre devrait correspondre à ce qui se passe en cas d'emphysème (voir Section 2.2.3). Au contraire, une augmentation de k correspondrait à un arbre bronchique plus rigide, ce qui arrive en cas de fibrose par exemple. Ensuite, nous modifions la valeur de θ pour $R_i = R_{out}$, et $k_0 = k_{00}$ fixés. θ mesure l'activité du muscle lisse ; un θ grand, correspond à une forte dilatation des bronches, ce qui peut arriver par exemple lors de l'utilisation d'un bronchodilatateur (par exemple la Ventoline). Enfin, nous modifions les résistances R_i aux sorties Γ_i, avec $k_0 = k_{00}$ et $\theta = 0.011$, ce qui pourrait modéliser une tumeur au niveau des sous–arbres correspondants. Dans ce cas, nous traçons la norme de la vitesse du fluide et la pression (voir Figure 3.12) au temps $t = 20$s. Ces graphes doivent être mis en regard avec ceux correspondant au cas précédent, où toutes les résistances étaient égales (Figure 3.9). La Figure 3.13 montre l'influence des modifications des résistances sur le portrait de phase.

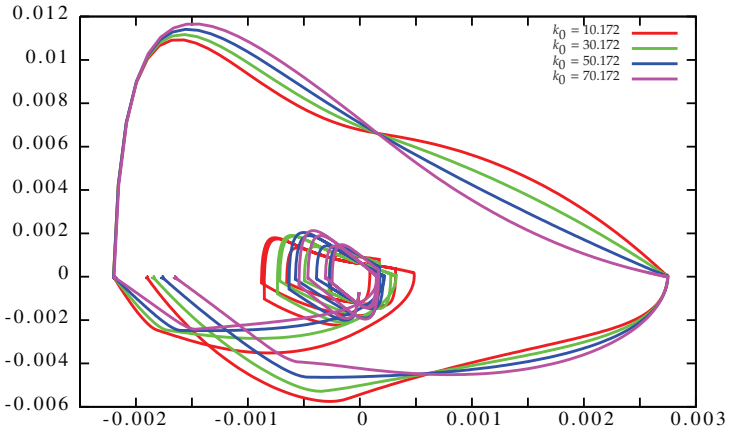


FIG. 3.7: Portrait de phase flux (en $m^3 \cdot s^{-1}$) – volume (en m^3) pour différentes valeurs de k_0 autour de la valeur physiologique $k_{00} = 40.172$

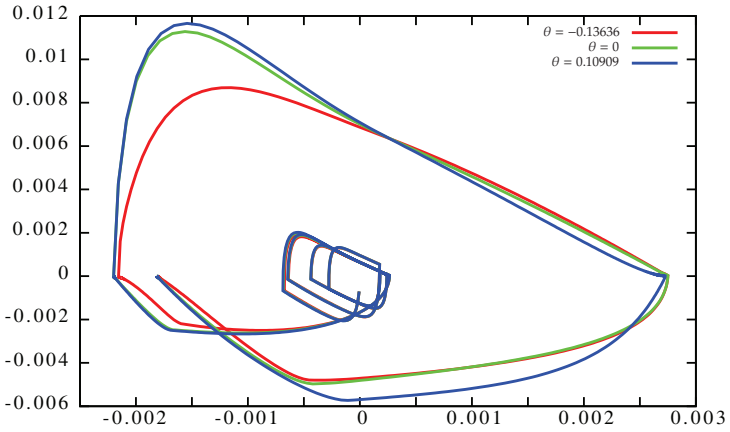
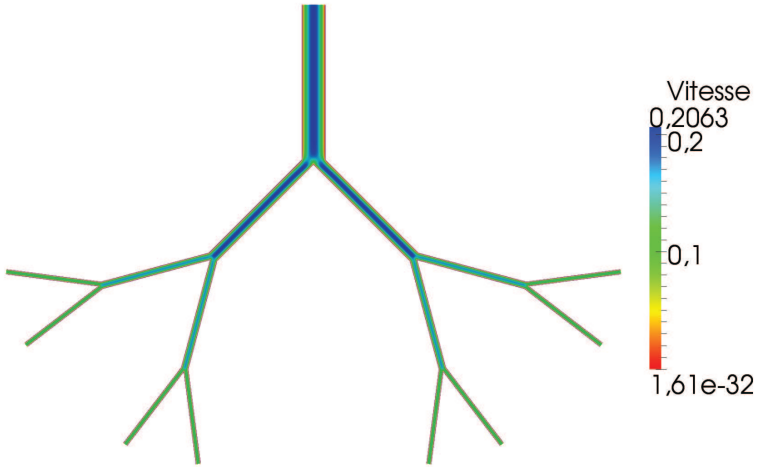
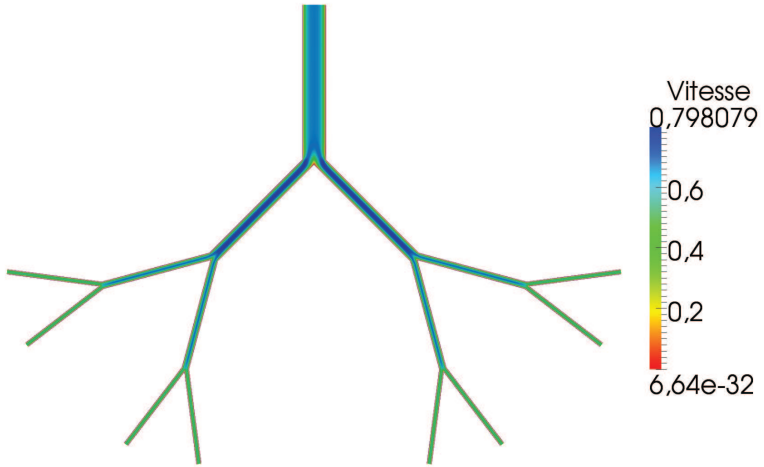


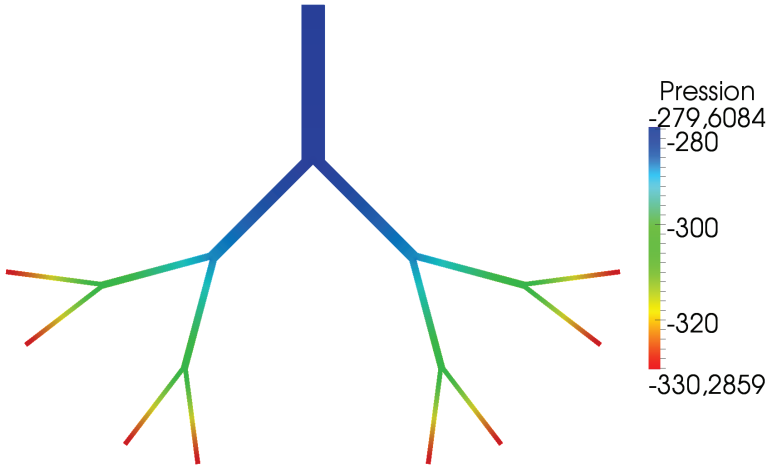
FIG. 3.8: Portrait de phase flux (en $m^3 \cdot s^{-1}$) – volume (en m^3) pour différentes valeurs de θ entre $\theta_{min} = -0.13636$ et $\theta_{max} = 0.10909$ (donnés par x_{min}, x_{max} et la formule (3.33)).



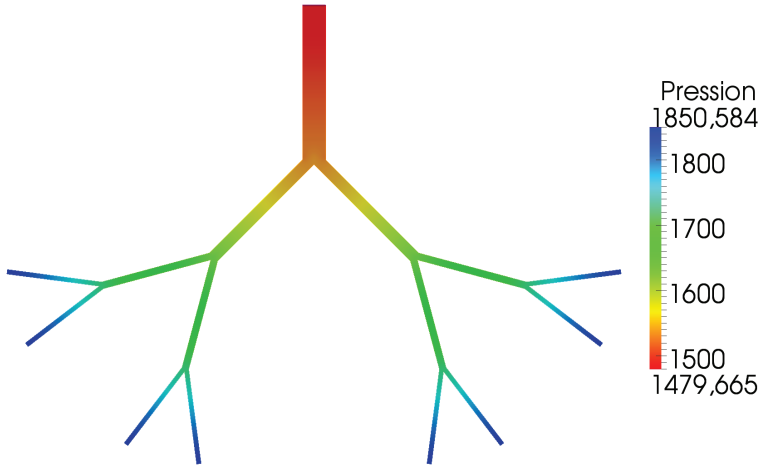
(a) Vitesse à l'inspiration maximale $t = 9\,\mathrm{s}$



(b) Vitesse à l'expiration maximale $t = 10.5\,\mathrm{s}$

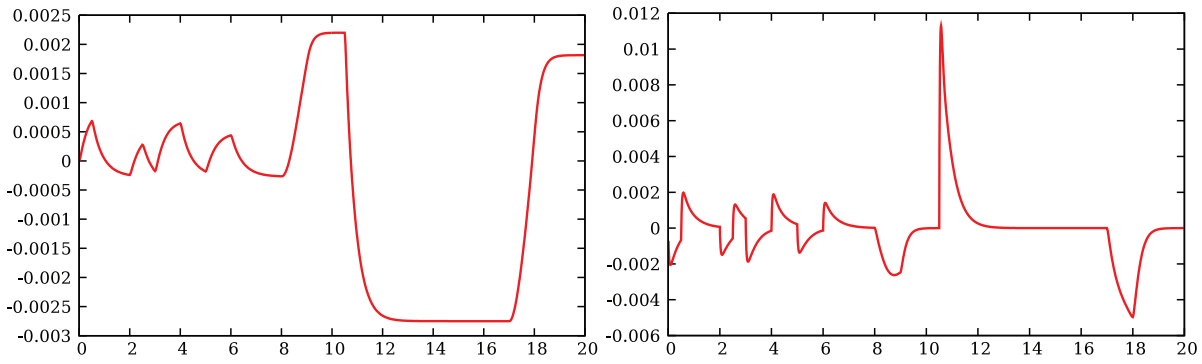


(c) Pression à l'inspiration maximale $t = 9\,\mathrm{s}$



(d) Pression à l'expiration maximale $t = 10.5\,\mathrm{s}$

FIG. 3.9: Vitesse de l'air (en $\mathrm{m} \cdot \mathrm{s}^{-1}$) et pression (en Pa) pour des données "moyennes" dans le poumon aux temps $t = 9$ et $t = 10.5\,\mathrm{s}$, *i.e.* à inspiration et expiration maximales (voir Figure 3.4). Visualisation à l'aide du logiciel `Paraview`.



(a) Volume (b) Flux

FIG. 3.10: Variations de volume (en m^3) et flux d'air (en $m^3 \cdot s^{-1}$) à travers le bord Γ_0 en fonction du temps en régime forcé avec les données physiologiques correspondant à un patient normal.

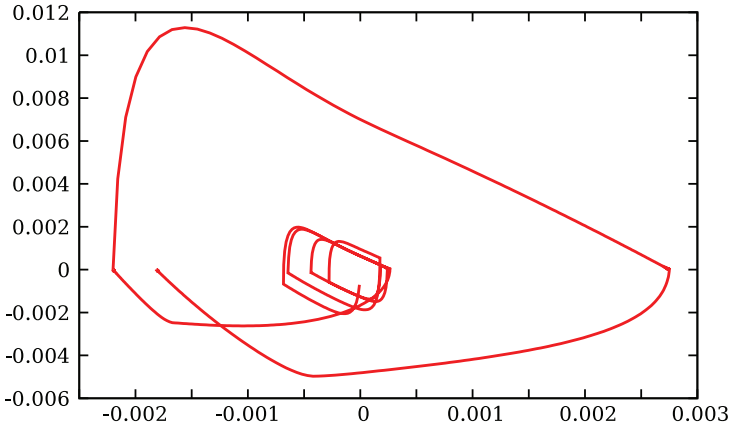
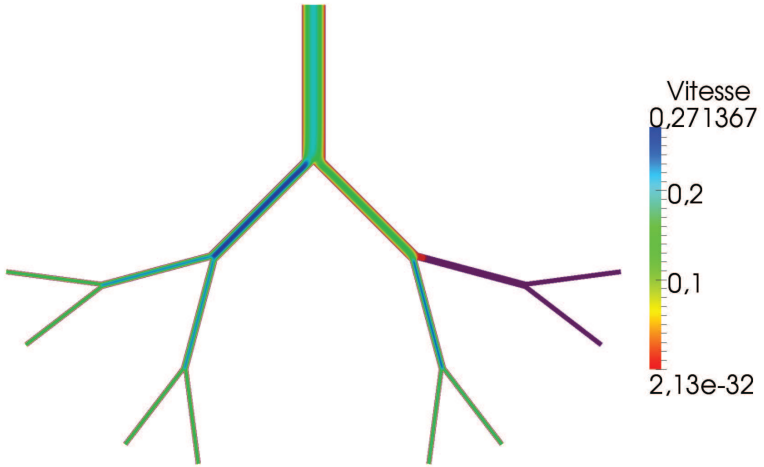
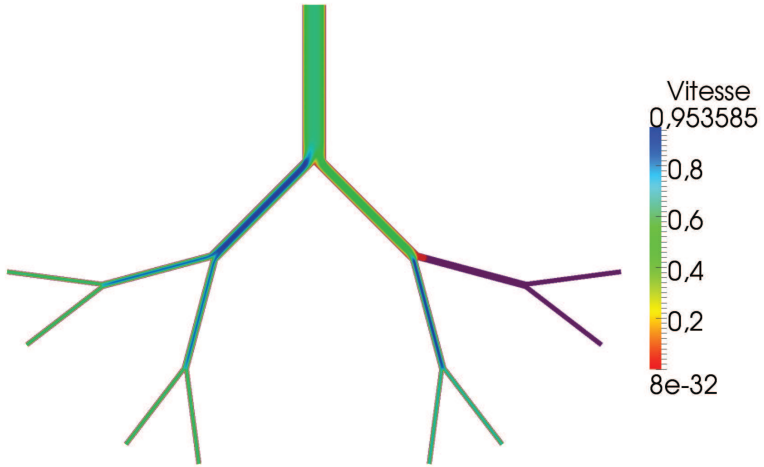


FIG. 3.11: Portrait de phase flux (en $m^3 \cdot s^{-1}$) – volume (en m^3).

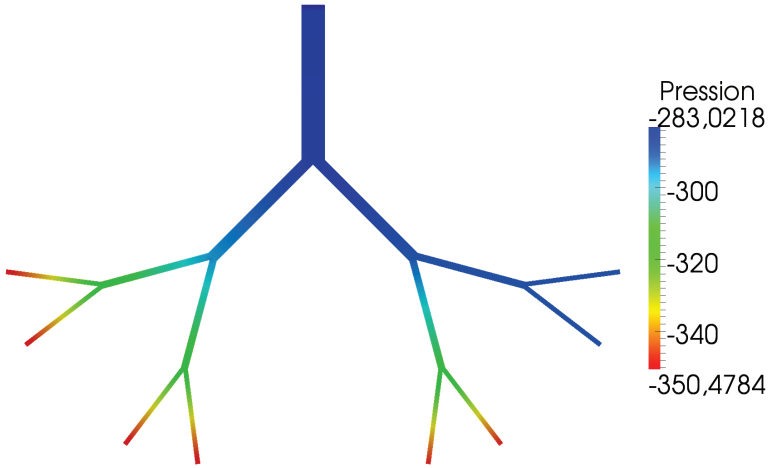
Qualitativement, nous retrouvons les courbes obtenues par spirométrie. Pour s'en convaincre comparer la Figure 3.11 au résultat d'examen médical Figure 2.6 ou bien au résultat d'examen non pathologique donné dans [MMSS08]. Insistons sur le fait que nous retrouvons les bons ordres de grandeur pour les volumes et débits.



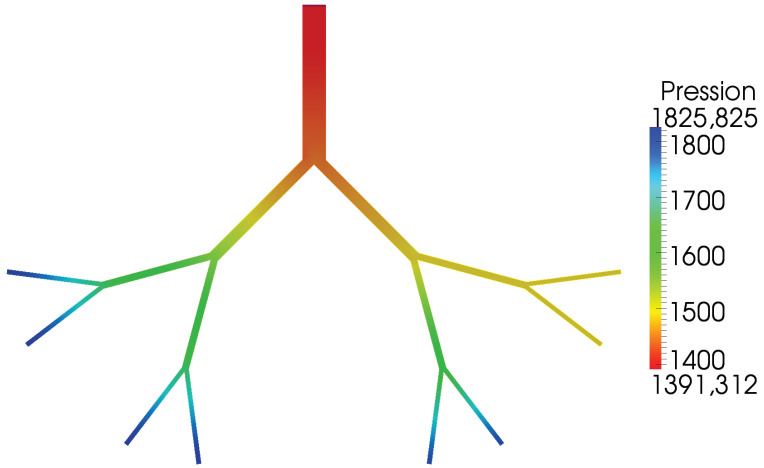
(a) Vitesse à l'inspiration maximale $t = 9\,\mathrm{s}$



(b) Vitesse à l'expiration maximale $t = 10.5\,\mathrm{s}$



(c) Pression à l'inspiration maximale $t = 9\,\mathrm{s}$



(d) Pression à l'expiration maximale $t = 10.5\,\mathrm{s}$

FIG. 3.12: Vitesse de l'air (en $\mathrm{m} \cdot \mathrm{s}^{-1}$) et pression (en Pa) dans le poumon aux temps $t = 9$ et $t = 10.5\,\mathrm{s}$, *i.e.* à inspiration et expiration maximales (voir Figure 3.4), pour $i = 1..6$, $R_i = R_{out}$ sur Γ_i (voir (3.33)) et sur Γ_7 et Γ_8, $R_i = R_{out} \times 10^6$. Visualisation à l'aide du logiciel `Paraview`.

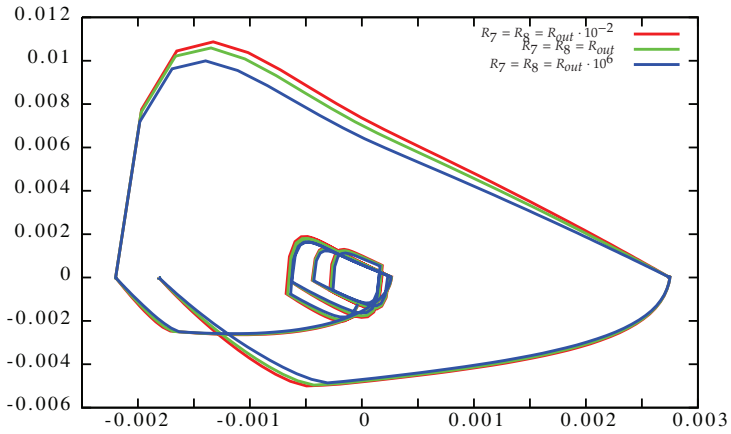


FIG. 3.13: Portrait de phase flux (en $m^3 \cdot s^{-1}$) – volume (en m^3) pour différentes valeurs de R_a (voir (3.33)) les résistances aux sorties Γ_7 et Γ_8. Pour $i = 1, \ldots, 6$, $R_i = R_{out}$ sur Γ_i et sur Γ_7 et Γ_8, $R_i = R_{out} \times 10^{2(s-1)}$.

Commentaires :

- La sensitivité du portrait de phase du modèle couplé semble plus importante vis–à–vis de la constante de raideur du ressort k et du paramètre θ que vis–à–vis des résistances R_i.
- Néanmoins, nous observons dans les branches correspondant aux sorties où l'on a augmenté la résistance R_i, une diminution du flux.
- La Figure 3.13, représente le portrait de phase dans le cas où nous avons augmenté la résistance sur deux sorties, Γ_7 et Γ_8. On observe que le débit de pointe (DEP) (voir Figure 2.4) décroît lorsque la résistance augmente.
- Pour k_0 grand, à la Figure 3.13, nous observons que la courbe d'expiration (partie supérieure, voir Figure 2.4), se creuse. Ceci, correspondrait plutôt au cas d'un syndrome obstructif (voir Figure 2.5).
- Dans notre modèle non linéaire, tel qu'il est écrit, on ne peut pas observer de syndrome restrictif (voir Figure 2.5). En effet, étant donné la loi choisie pour la dépendance de la constante de raideur k (voir (2.11), page 84) par rapport à x, toutes nos $(S\,x_{min}, 0)$ courbes (Figures 3.11, 3.7, 3.8 et 3.13) passent par les points de coordonnées $(S\,x_{max}, 0)$. Si on suppose uniquement $k = k_0$ constante pour différents choix de k_0, on peut retrouver qualitativement le syndrome restrictif lorsque $k_0 < k_{00}$, caractéristique de la fibrose, ainsi qu'une capcité totale plus grande dans le cas $k_0 > k_{00}$, caractéristique d'un emphysème par exemple. Nous traçons à la Figure 3.14 les différents portraits de phase obtenus pour différentes valeurs de k, supposé constant pour ces simulations. Nous observons que le débit expiratoire pointe (DEP), c'est–à–dire la maximum du pic d'expiration (voir Section 2.2.3, page 76) varie peu. D'après les médecins, dans le cas d'un emphysème, on observe également une diminution du DEP, ce que nous ne retrouvons pas ici. Sans doute, faudrait-il modifier également le paramètre θ pour retrouver exactement toutes les caractéristiques de l'emphysème.

3.4.5 Simulations 3D

Les simulations 2D ne peuvent pas reproduire l'ensemble des effets 3D, notamment les ajustements sur les paramètres ne permettent pas de calculer la valeur exacte de la norme de la vitesse (voir page 105). Des simulations 3D doivent donc être réalisées et exploitées. La méthode numérique utilisée peut tout à fait être appliquée dans le cas d'une géométrie 3D. Celle–ci a été implémentée par DRISS YAKOUBI sur un domaine obtenu par segmentation à partir de l'imagerie médicale (voir Figure 3.15). Il s'agit de la même géométrie 3D utilisée dans [FMP+05].

Pour les simulations 3D, ce sont donc les paramètres réels donnés à la Table 5.2 que nous utilisons.

Nous ne rentrons pas ici dans les détails du calcul 3D ; toutefois, nous donnons quelques résultats 3D afin de montrer que les comportements obtenus en 2D sont en tout point semblables à ceux obtenus en 2D.

La Figure 3.16 représente la vitesse et la pression à l'inspiration maximale $t = 9\,s$ et expiration maximale $t = 10.5\,s$. La Figure 3.17 représente le portrait de phase calculé avec la géométrie 3D. À nouveau, nous obtenons

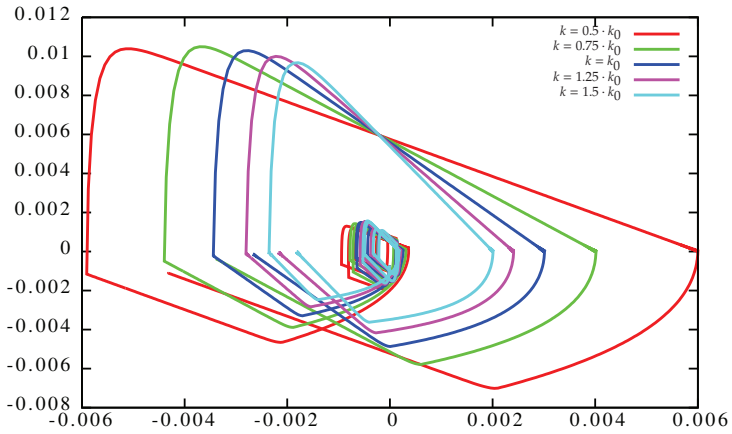


FIG. 3.14: Portrait de phase flux (en $m^3 \cdot s^{-1}$) – volume (en m^3) pour k supposé constant au cours de la manoeuvre. Nous traçons la courbe pour 5 différentes valeurs de k : $k = 0.5\,k_0$, $0.75\,k_0$, k_0, $1.25\,k_0$, $1.5\,k_0$

une courbe comparable à celles obtenues par examen médical, et nous notons que les ordres de grandeurs sont cohérents avec les mesures expérimentales.

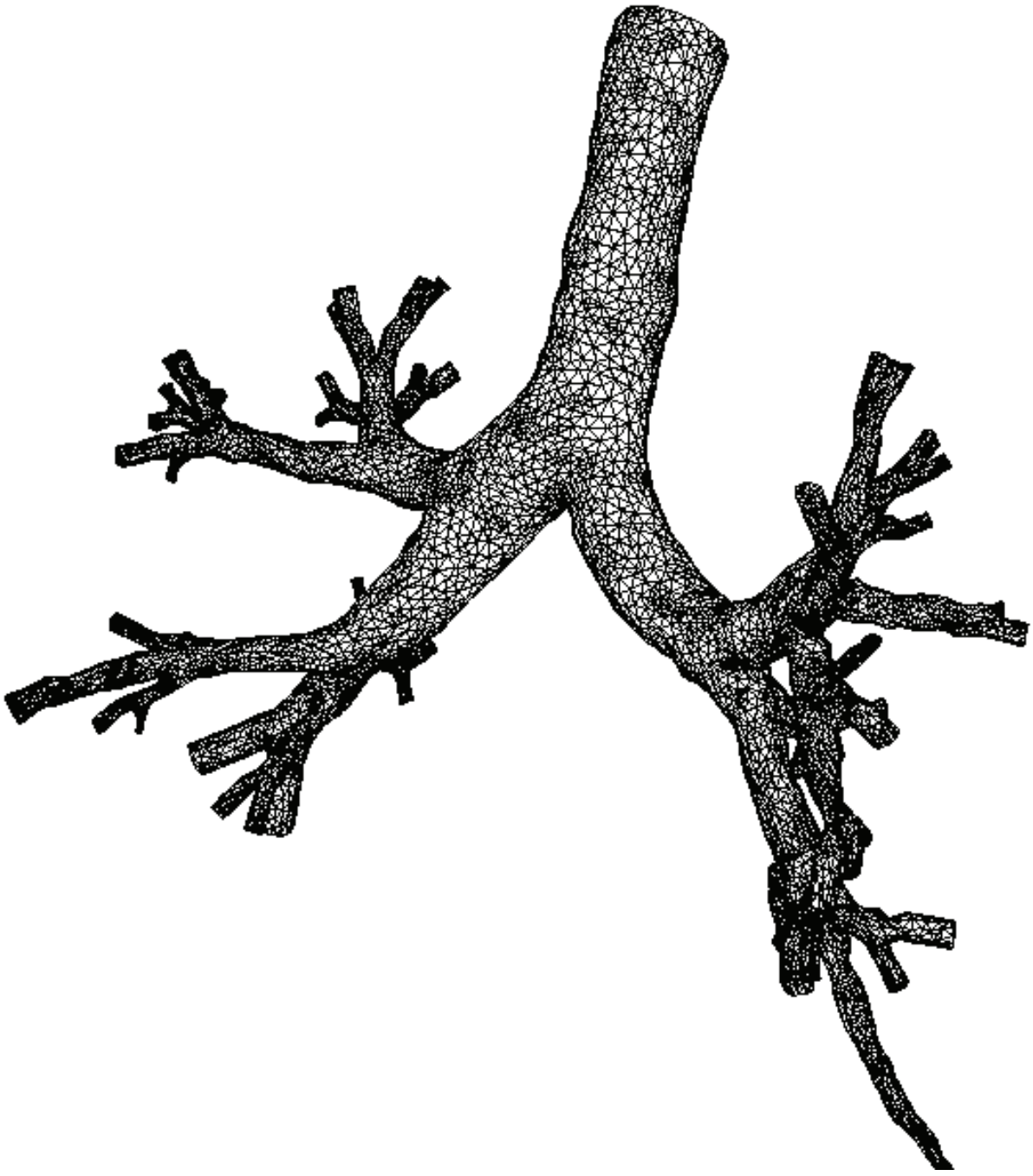
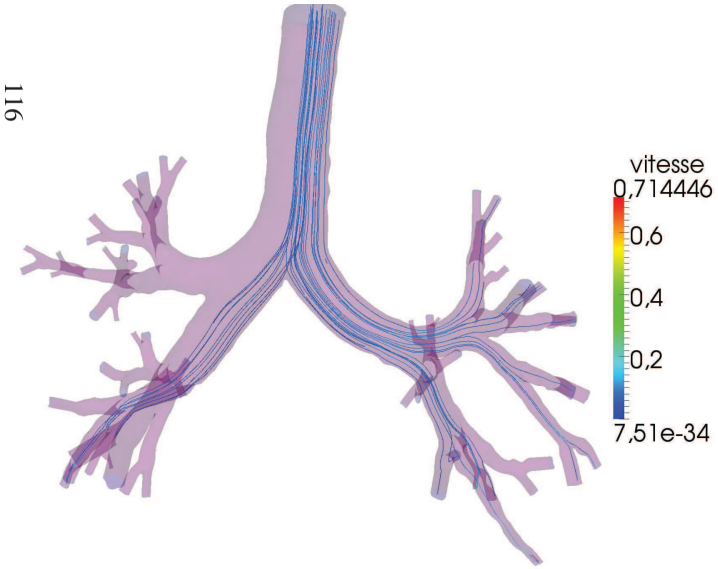
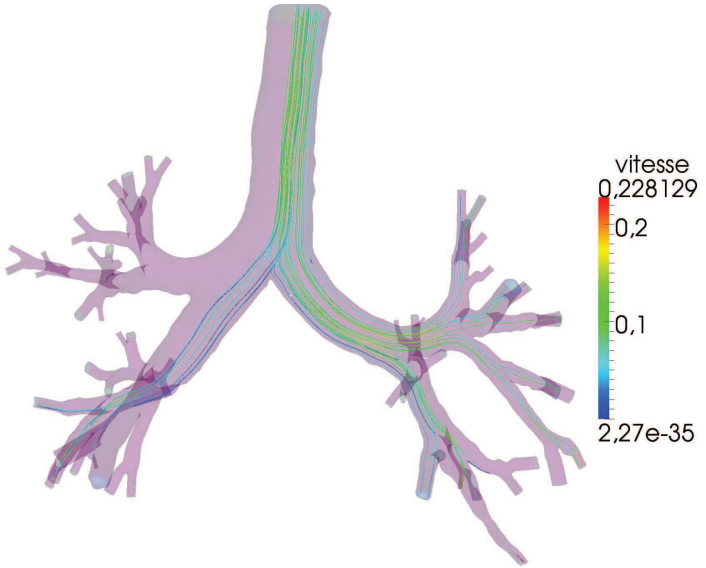


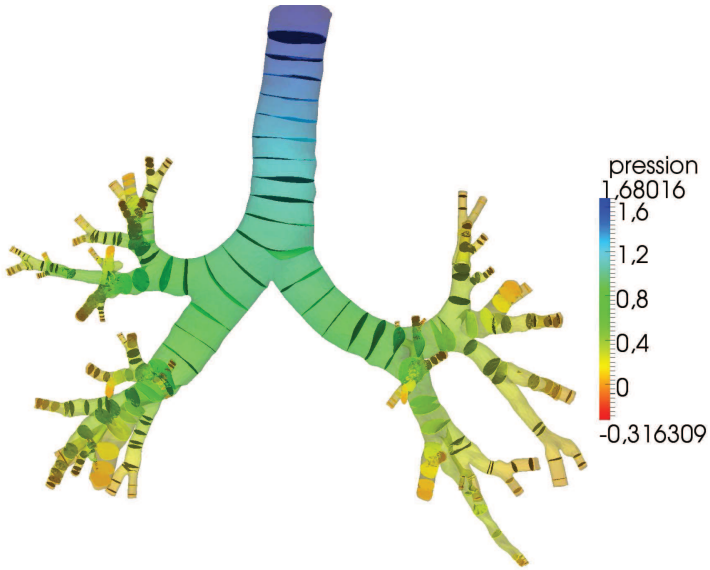
FIG. 3.15: Maillage utilisé pour le calcul 3D



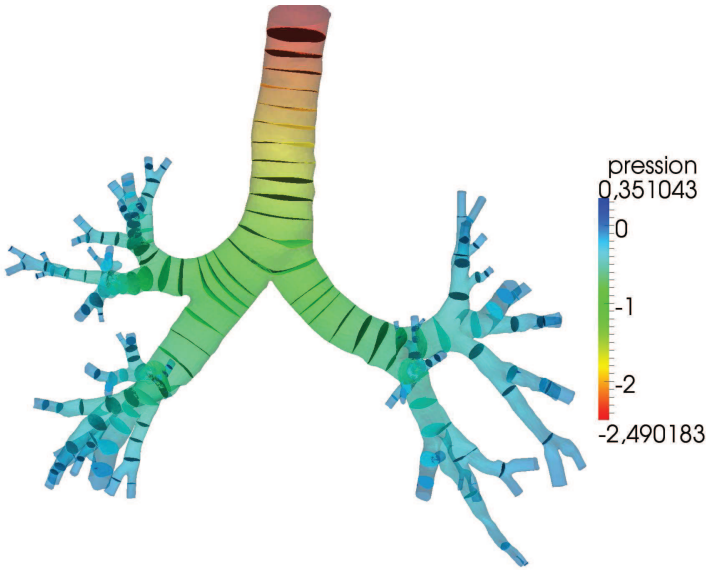
(a) Vitesse à l'inspiration maximale $t = 9\,\mathrm{s}$



(b) Vitesse à l'expiration maximale $t = 10.5\,\mathrm{s}$



(c) Pression à l'inspiration maximale $t = 9\,\mathrm{s}$



(d) Pression à l'expiration maximale $t = 10.5\,\mathrm{s}$

FIG. 3.16: Vitesse de l'air (en $\mathrm{m} \cdot \mathrm{s}^{-1}$) et pression (en Pa) dans le poumon aux temps $t = 9$ et $t = 10.5\,\mathrm{s}$, *i.e.* à inspiration et expiration maximales (voir Figure 3.4), pour une géométrie 3D. Visualisation à l'aide du logiciel `Paraview`.

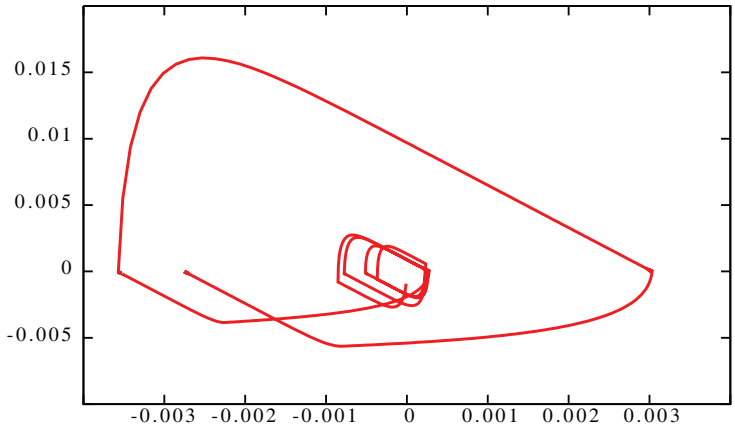


FIG. 3.17: Portrait de phase flux (en $\mathrm{m}^3 \cdot \mathrm{s}^{-1}$) – volume (en m^3) pour la géométrie 3D.

3.5 Conclusions et perspectives

D'un point de vue numérique, nous avons développé une méthode facilement implémentable dans tout solveur numérique. Il y a cependant un inconvénient majeur à notre méthode : en effet, nous ne pouvons traiter le terme convectif que par une méthode explicite ou bien par les caractéristiques (ce qui a été choisi ici). Toutefois, complétée par une méthode de projection pour la résolution du problème de Stokes, cette méthode nous permet d'obtenir des résultats convenables en un temps raisonnable.

D'un point de vue biologique, nous obtenons des portraits de phase comparables à ceux fournis par des examens médicaux de spirométrie (voir [MMSS08]). La comparaison des résultats pour différentes valeurs de θ, k, et R_i sont encourageantes : ce modèle pourra sans doute aider à une meilleure compréhension de certaines pathologies bronchiques telles l'asthme (modification du θ) ou l'emphysème (modification de k).

D'un point de vue de modélisation, le modèle peut encore être amélioré ; par exemple, d'autres modèles élastiques non–linéaires devraient sans doute être développés afin de mieux capter la complexité du phénomène, par exemple pour retrouver les caractéristiques d'un syndrome restrictif (voir Figure 2.5). En particulier la condition qui impose le volume maximal et le volume minimal, conséquence de la forme particulière choisie pour décrire la constante de raideur au cours du temps (voir (2.11)) semble trop restrictive et ne permet pas de décrire certaines pathologies, notamment les pathologies dites obstructives.

Remerciements

Mes co–auteurs et moi–même souhaitons remercier MATTEO ASTORINO, LAURENT BOUDIN et MOURAD ISMAIL pour les discussions très fructueuses que nous avons eues au cours et à la suite du CEMRACS 2008.

Merci beaucoup à Bérénice et Céline pour leurs relectures attentives (et estivale !) et leurs remarques sur ce chapitre.

Chapitre 4

Un modèle pour l'absorption de l'oxygène dans les poumons

Dans ce chapitre, nous nous intéressons à la modélisation du système respiratoire, en particulier des échanges gazeux dans l'arbre bronchique. Notre but est de coupler le modèle mécanique simple décrit à la Section 2.4.1 à un modèle précis de transport d'oxygène, permettant d'intégrer les effets d'hétérogénéité des échanges gazeux le long de l'arbre bronchique.
Dans cette perspective, deux modèles ont été élaborés. Le premier est un modèle monocompartiment du système respiratoire permettant de prévoir le comportement du système durant une situation dynamique complexe. Le second est un modèle d'absorption de l'oxygène qui prend en compte l'hétérogénéité des transferts gazeux en fonction de la profondeur dans l'arbre bronchique.
Le modèle parvient à reproduire le comportement mécanique du poumon ainsi que les données observables liées aux échanges d'oxygène, telles que la quantité moyenne d'oxygène transférée dans le sang pendant un cycle respiratoire, dans les deux régimes que nous avons étudiés : respiration normale et respiration à l'exercice. Le modèle confirme également qu'un arbre bronchique parfaitement rigide serait contre-productif en empêchant des quantités suffisantes d'oxygène d'atteindre les alvéoles, comme cela a été montré par MARTIN, SIMILOWSKI, STRAUS & MAURY [MMSS08].
Ce travail fait l'objet d'une collaboration avec SÉBASTIEN MARTIN et BERTRAND MAURY. C'est un travail encore en cours.

4.1 Introduction

Un grand nombre de modèles mathématiques pour le système respiratoire ont déjà été proposés (voir [MMSS08] et les références qu'il contient pour une rapide description des modèles existants). Dans ces modèles, le poumon est souvent représenté par un système multi–compartiments muni de conditions de couplage particulières. En général, le niveau de détail est fortement lié au nombre de paramètres. Dans le travail qui suit, nous préférerons encore un modèle le plus simple possible vis–à–vis des quantités mesurables par les cliniciens, afin de nous concentrer sur la compréhension de l'influence des caractéristiques principales de la convection bronchique sur les échanges gazeux. Notre travail est largement inspiré de [MMSS08] ; cependant dans ce travail les auteurs considèrent uniquement la concentration moyenne en oxygène dans le poumon. Ce cadre de travail n'est pas suffisant pour une description détaillée de la distribution du gaz le long de l'arbre bronchique (en particulier, il ne permet pas de décrire l'appauvrissement en oxygène le long du parcours bronchique). Nous avons vu (voir Section 2.2.4, page 78) que les alvéoles, *i.e.* les unités d'échange, sont disposées le long de l'arbre bronchique à partir de la 15e génération.

Ainsi la ventilation des alvéoles se fait en série le long des voies respiratoires. L'efficacité de l'échange gazeux est tributaire de l'apport en oxygène dans l'alvéole et donc de sa proximité de l'entrée de l'acinus ou à sa périphérie. Rappelons encore que dû à la géométrie dichotomique de l'arbre bronchique et du facteur d'homothétie, la moitié de la surface d'échange totale disponible est située sur la dernière génération.

Soulignons maintenant deux problèmes qui affectent l'efficacité des échanges gazeux et que nous voulons prendre en compte dans notre modèle :

1. Puisque les voies respiratoires ont la forme d'un arbre dichotomique, leur section totale augmente exponentiellement avec les générations (voir Section 2.2.2). En conséquence, la vitesse de l'air décroît le long de l'arbre et devient à partir d'un certain point plus petite que la vitesse de diffusion de l'oxygène dans l'air. Cette transition apparait grossièrement à l'entrée de l'acinus, mais peut être décalée selon le régime respiratoire auquel on se place.
2. Tandis que l'oxygène diffuse le long des voies acinaires, il est progressivement absorbé à travers la surface d'échange développée le long de sa route. Se faisant, la pression partielle en oxygène décroît. Ce processus, souvent appelé effets d'écrantage (*screening* en anglais), est important car il détermine quelle quantité d'oxygène atteint les parties les plus profondes du poumon, où se concentre la plus grande partie de la surface d'échange. Il y a compétition entre deux phénomènes : d'une part, dans les parties hautes de l'acinus, c'est–à–dire proche de l'entrée, la pression partielle en oxygène est élevée tandis que la surface d'échange disponible est petite. D'autre part, dans les parties plus périphériques, la pression partielle en oxygène est faible, suite aux effets d'écrantage et à l'appauvrissement en oxygène, mais la surface d'échange est grande. Un raisonnement naïf permetterait de dire que pour annuler ces effets, il suffirait de respirer plus fort ; toutefois la notion de rendement est ici pertinente, l'énergie fournie pour la respiration serait alors trop importante en comparaison avec l'énergie fournie pour l'activité de l'individu. À l'effort, la notion de rendement est moins pertinente ; le sujet cherche à absorbé un maximun d'oxygène afin d'alimenter les muscles.

Dans [GH06], l'efficacité de la membrane pulmonaire lors des échanges gazeux entre les alvéoles et le sang pulmonaire est décrite quantitativement en termes de *capacité de diffusion*, qui est défini comme le volume de gaz qui diffuse à travers la membrane par unité de temps pour une différence de pression de 1 mmHg. Au repos, pour un jeune homme moyen, la capacité de diffusion de l'oxygène est d'environ $0.35 \cdot 10^{-6}\ \mathrm{m}^3 \cdot \mathrm{s}^{-1} \cdot \mathrm{mm\,Hg}^{-1}$. En termes de quantité d'oxygène absorbée, étant donné que la différence de pression partielle entre les deux côtés de la membrane est de l'ordre de $11\,\mathrm{mm\,Hg}$, nous déduisons le flux d'oxygène diffusant à travers la membrane est de l'ordre de $3.8 \cdot 10^{-6}\ \mathrm{m}^3 \cdot \mathrm{s}^{-1}$. Cette valeur est cohérente avec les données dans [Wei84]. Que se passe-t-il à l'exercice ? Premièrement, puisque la force à l'inspiration est plus importante, le transport convectif de l'air conduit à une meilleure ventilation des alvéoles avec de l'air frais. Deuxièmement, l'apport sanguin au poumon augmente et rythme cardiaque et rythme ventilatoire se rapprochent, *i.e.* le ratio ventilation/perfusion se rapproche de 1. GUYTON *et al.* [GH06] estiment qu'à l'exercice, pour un jeune homme sain, la capacité de diffusion augmente à $1.08 \cdot 10^{-6}\ \mathrm{m}^3 \cdot \mathrm{s}^{-1} \cdot \mathrm{mmHg}^{-1}$.

Notre objectif est d'expliquer cette augmentation de la capacité de diffusion. Nous verrons que la géométrie particulière du poumon est en grande partie responsable de cette différence de comportement entre le repos et l'exercice.

Nous procéderons comme suit ; afin de comprendre un peu mieux les quantités décrites dans [GH06], nous retrouverons ces résultats pour un modèle 0D pour la diffusion de l'oxygène, couplé au modèle mécanique déjà décrit précédemment (Section 2.4.1, page 83). Ce modèle mono–compartiment ne sera bien entendu pas suffisant pour décrire l'hétérogénéité de l'absorption de l'oxygène le long de l'arbre bronchique qui, nous le verrons, semble essentielle pour comprendre le changement de régime repos/exercice. Nous nous tournerons donc vers un modèle 1D pour l'oxygène, où la dimension d'espace correspondra au chemin parcouru le long de l'arbre bronchique.

Avant de rentrer dans les détails, rappelons brièvement les deux principales raisons physiologiques pouvant expliquer cette différence entre les deux régimes.

1. Comme déjà mentionné, nous observons à l'exercice une meilleure correspondance entre rythme cardiaque et rythme respiratoire ; cela est mesuré par le *ratio ventilation/perfusion* noté $\rho_{V/P}$. Nous trouvons dans [GH06] qu'au repos nous avons plus ou moins $\rho_{V/P}^{\mathrm{repos}} = 1/6$; cela signifie qu'au repos durant un cycle respiratoire le coeur bat en moyenne 6 fois. A l'exercice ce ratio augmente, il est estimé à $\rho_{V/P}^{\mathrm{effort}} = 1/2$.
2. Lorsque on passe du régime de repos à celui de l'effort, beaucoup d'alvéoles pulmonaires auparavant inutilisées s'ouvrent et d'autres déjà ouvertes subissent une dilatation supplémentaire, accroissant ainsi la surface

d'échange. C'est pourquoi, dans notre modèle nous devons prendre en compte la variation de la surface d'échange dans la mesure où la captation de l'oxygène est propre à la surface d'échange. En d'autres mots, nous devons prendre en compte les effets d'écrantage.

En résumé, le flux d'oxygène à travers la membrane pulmonaire doit être proportionnel à :

- la différence des pressions partielles en oxygène de chaque coté de la membrane, à savoir entre l'alvéole et le sang pulmonaire ;
- la surface alvéolaire disponible ;
- la capacité de diffusion de la membrane qui est constante ;
- le ratio ventilation/perfusion.

Validation de notre modèle

Nous voulons construire un modèle précis pour le transport de l'oxygène dans l'arbre bronchique permettant de mieux comprendre l'hétérogénéité du transfert de l'oxygène dans le sang. La validation de ce modèle repose sur deux critères principaux :

1. quantitativement, les volumes moyens d'oxygène absorbés par unité de temps au repos et à l'exercice doivent correspondre à ceux que l'on peut trouver dans la littérature. Au repos, ce flux moyen est égal à $4 \cdot 10^{-6}\ \mathrm{m}^3 \cdot \mathrm{s}^{-1}$ [Wei84, GH06]. Concernant le volume d'oxygène absorbé par unité de temps à l'effort, la littérature est moins univoque : nous trouvons des valeurs allant de $35 \cdot 10^{-6}\ \mathrm{m}^3 \cdot \mathrm{s}^{-1}$ [Wei84] à $54 \cdot 10^{-6}\ \mathrm{m}^3 \cdot \mathrm{s}^{-1}$ [SFW02].
2. Nous voulons observer les effets d'écrantage au repos ; à savoir au repos la plupart de l'oxygène absorbé doit l'être via les alvéoles situées sur les premières générations (entre la 15e et la 18e), tandis que à l'exercice, puisque les alvéoles sont bien mieux ventilées, l'oxygène est principalement absorbé par les dernières générations. [WSF05].

4.2 Un modèle mono–compartiment pour l'oxygène

Les données obtenues par diverses expériences suggèrent toutes l'existence d'interactions complexes entre le système respiratoire et le système cardio–vasculaire chez l'homme. Ces interactions comprennent notamment la synchronisation des rythmes respiratoire et cardiaque. Des modèles mathématiques ont été développés afin de décrire les mécanismes sous–jacents et améliorer notre compréhension de ce phénomène physiologique. Historiquement, les échanges gazeux ont été étudiés à différents niveaux de détails. Puisque la fonction du poumon est d'apporter l'oxygène de l'atmosphère extérieur vers le sang, l'idée d'un tel modèle est de reproduire les connections entre l'extérieur et les alvéoles.

Dans cette partie préliminaire, nous proposons de coupler le modèle mécanique 0D déjà décrit à la Section 2.4.1 avec un modèle d'échange gazeux 0D également.

Rappelons rapidement les équations et les notations ; pour la partie mécanique, le poumon est considéré comme une boîte avec une paroi mobile reliée à un ressort, modélisant le parenchyme, voir Figure 2.8. Nous notons m la masse déplacée, S la surface de la paroi mobile et x le déplacement tel que la variation de volume pulmonaire soit donnée par $\Delta V = Sx$. La constante μ représente la résistance des tissus, $R(x)$ la résistance de l'arbre bronchique, qui dépend du déplacement ainsi que $k(x)$ la constante de raideur du ressort, voir page 84 pour plus de détails. L'équation mécanique obtenue (voir Section 2.4.1, page 83) est la suivante :

$$m\ddot{x} + (R(x)S^2 + \mu)\dot{x} + k(x)x = f_{ext},$$

où f_{ext} désigne les forces d'inspiration ou d'expiration contrôlées par le muscle diaphragmatique. Il s'agit de la quantité contrôlée par le sujet afin de piloter le mécanisme de ventilation.

La fonction principale du poumon est le transport de l'oxygène depuis l'atmosphère vers le sang, et d'extraire le dioxyde de carbone du sang vers l'atmosphère : ainsi, l'efficacité de la ventilation ne doit pas reposer sur la quantité d'air entrant dans le poumon mais plutôt sur la quantité d'oxygène transportée jusqu'au sang. Le flux d'O_2, cependant corrélé au flux d'air entrant, est également lié à la taille de la surface d'échange. Cette surface (sous des

hypothèses de déformations homothétiques), notée S_A dans la suite, est décrite par une fonction dépendant du volume total des alvéoles, *i.e.*

$$S_A := V_A^{2/3} = \left(V_A^0 + (1-\theta)Sx\right)^{2/3}$$

Nous noterons P_A la pression partielle en oxygène dans les alvéoles et P_{sang} celle dans le sang, lorsque celui-ci entre dans le réseau vasculaire pulmonaire. Nous supposons également que le flux d'oxygène est proportionnel à la différence des pressions partielles entre l'alvéole et le sang et au ratio ventilation/perfusion. Ainsi, nous obtenons :

$$\dot{V}_{O_2} = \beta^* \, \rho_{V/P} \, V_A^{2/3} \, (P_A - P_{\text{sang}}),$$

où $\rho_{V/P}$ est le ratio ventilation/perfusion et β^* est une constante représentant la capacité de diffusion de la membrane pour l'oxygène (voir [GH06] p.498). Comme mentionné à la Section 2.2.4, le paramètre $\rho_{V/P}$ est donné dans la littérature [GH06] : au repos $\rho_{V/P} = 1/6$ et à l'exercice $\rho_{V/P} = 1/2$. Nous trouvons également dans la littérature $P_{\text{sang}} = 40\,\text{mmHg}$. Si c_A représente la fraction molaire en oxygène dans les alvéoles, nous avons la relation suivante : $P_A = P_0 \, c_A$, où P_0 est la pression atmosphérique ($P_0 = 760\,\text{mmHg}$).

Afin de compléter notre modèle, nous écrivons le bilan de masse pour l'oxygène dans les alvéoles. La quantité d'oxygène dans les alvéoles est sujette aux phénomènes suivants :

i) Les alvéoles sont fournis en air frais via l'arbre bronchique ; ce flux d'oxygène est modélisés par le produit de $\dot{V}_A$ avec la concentration en oxygène à la frontière ouverte du domaine, c'est–à–dire à l'entrée des alvéoles, qui est égale à c_0 (concentration en oxygène de l'air frais, $c_0 = 0.20$) à l'inspiration ou bien c_A à l'expiration. Rappelons la notation :

$$[y \geq 0\,?\,a, b] = \begin{cases} a & \text{si } y \geq 0, \\ b & \text{sinon.} \end{cases}$$

Ainsi, le flux d'oxygène de l'extérieur vers les alvéoles s'écrit $\dot{V}_A\,[\dot{x} \geq 0\,?\,c_0, c_A]$.

ii) L'oxygène diffuse à travers la membrane alvéolaire vers le sang. Comme mentionné plus haut, le flux correspondant s'écrit $\dot{V}_{O_2}$.

Finalement, la variation totale de la quantité d'oxygène dans les alvéoles vérifie :

$$\dot{c}_A = \frac{1}{V_A}\left(\dot{V}_A\,[\dot{x} \geq 0\,?\,c_0 - c_A, 0] - \dot{V}_{O_2}\right).$$

En couplant ce modèle d'échange gazeux avec le modèle mécanique, nous obtenons le modèle complet, qui peut être réécrit sous la forme d'un système d'équations différentielles ordinaires du premier ordre :

$$\begin{cases} \dot{x} &= u, \\ \dot{u} &= \dfrac{1}{m}\left(f_{ext} - \left(\dfrac{R_0 S^2}{1 + \theta S x / V_B^0} + \mu\right) u - kx\right), \\ \dot{V}_{O_2} &= \beta^* \, \rho_{V/P} \left(V_A^0 + (1-\theta)Sx\right)^{2/3} \, (c_A \, P_0 - P_{\text{sang}}), \\ \dot{c}_A &= \dfrac{1}{V_A^0 + (1-\theta)Sx}\left((1-\theta)S\dot{x}\,[\dot{x} \geq 0\,?\,c_0 - c_A, 0] - \dot{V}_{O_2}\right). \end{cases} \tag{4.1}$$

La constante de proportionnalité a été intégrée dans le paramètre β^* et, en suivant [GH06], la capacité de diffusion de la membrane pulmonaire est égale à $D_{\text{poumon}} := 21\ \text{ml} \cdot \text{min}^{-1} \cdot \text{mmHg}^{-1}$ au repos ; à l'effort, pour un jeune homme sain la capacité de diffusion peut augmenter jusqu'à atteindre la valeur de $D_{\text{poumon}} := 65\ \text{ml} \cdot \text{min}^{-1} \cdot \text{mmHg}^{-1}$, qui correspond à trois fois la valeur au repos. Nous avons donc à comparer :

$$D_{\text{poumon}} = \beta^* \left(V_A^0\right)^{2/3},$$

ainsi nous obtenons :

$$\beta^* = \begin{cases} 1.65 \cdot 10^{-7}\,\text{m} \cdot \text{s}^{-1} \cdot \text{Pa}^{-1}, & \text{au repos,} \\ 5.12 \cdot 10^{-7}\,\text{m} \cdot \text{s}^{-1} \cdot \text{Pa}^{-1}, & \text{à l'exercice.} \end{cases}$$

En utilisant les paramètres donnés à la Table 4.1, les résultats numériques que nous obtenons sont assez précis pour prédire les quantités moyennes d'oxygène passées dans le sang dans les deux régimes ; ces résultats sont en bon accord avec les données expérimentales, voir Table 4.2.

pression atm. P_0	760	mmHg
pression partielle d'O_2 dans le sang P_{sang}	40	mmHg
ratio ventilation/perfusion $\rho_{V/P}$		
au repos	1/6	
à l'effort	1/2	
capacité de diffusion :		
au repos	$1.65 \cdot 10^{-7}$	$\text{m} \cdot \text{s}^{-1} \cdot \text{Pa}^{-1}$
à l'effort	$5.12 \cdot 10^{-7}$	$\text{m} \cdot \text{s}^{-1} \cdot \text{Pa}^{-1}$

TAB. 4.1: Données physiologique pour le modèle d'échange gazeux.

	GUYTON & HALL [GH06]	Modèle mathématique
au repos	$4 \cdot 10^{-6}$ m^3 s^{-1}	$5 \cdot 10^{-6}$ m^3 s^{-1}
à l'exercice	$35 \cdot 10^{-6}$ m^3 s^{-1}	$48 \cdot 10^{-6}$ m^3 s^{-1}

TAB. 4.2: Flux moyens d'oxygène vers le sang.

Commentaires

- Les valeurs obtenues par simulations numériques sont un peu sur–évaluées vis–à–vis des valeurs physiologiques (voir Table 4.2) ; cela est peut–être conséquence du fait que nous comptons deux fois l'augmentation de la surface efficace. En effet, la surface efficace croît via l'augmentation de volume et du facteur $\left(V_A^0 + (1-\theta)Sx\right)^{2/3}$ dans la troisième équation du système (4.1), mais cette augmentation de surface efficace est également prise en compte par l'augmentation de la capacité de diffusion lors du passage d'un régime de repos à un régime d'effort (voir Table 4.1).
- Un modèle mono–compartiment comme celui–ci ne permet pas de modéliser l'hétérogénéité de l'absorption de l'oxygène le long de l'arbre. Il permet cependant de retrouver les valeurs moyennes données dans la littérature de physiologie.
- Il n'est pas satisfaisant de choisir des valeurs différentes selon le régime pour la capacité de diffusion. Ce paramètre doit être caractéristique de l'épithélium pulmonaire, et constant quelque soit le régime respiratoire ou la localisation dans le poumon. Nous interprétons cette différence entre les deux valeurs pour la capacité de diffusion au repos et à l'effort donnée dans [GH06], comme une modification de la surface efficace, c'est–à–dire la surface atteinte par de l'air frais lors du cycle respiratoire. Dans l'idéal, nous aimerions que cette augmentation de la surface d'échange soit conséquence de la mécanique de la ventilation et de la géométrie particulière du poumon. Nous choisissons maintenant de parler de perméabilité notée W_{O_2}. En multipliant la perméabilité par la surface efficace, nous retrouvons ce que nous appelions jusqu'à présent la capacité de diffusion. W_{O_2} est maintenant une constante caractéristique de l'épithélium pulmonaire.
- Pour parler grossièrement, si on compare les volumes d'oxygène absorbés au repos et à l'exercice, nous devons gagner un facteur 9 (voir Table 5.4, page 138). Nous gagnons déjà un facteur 3 grâce à l'augmentation du ratio ventilation/perfusion (voir Table 5.5, page 138). Il nous manque encore un facteur 3. L'objectif de ce qui suit est de montrer que ce facteur 3 peut être obtenu grâce à la mécanique et à la géométrie.

4.3 Prise en compte de l'hétérogénéité : état de l'art

Le transfert de l'oxygène depuis l'air atmosphérique vers le sang se fait en trois étapes : la ventilation à travers les voies aériennes, la diffusion dans les sacs alvéolaires et finalement la diffusion à travers la membrane vers le sang pulmonaire. Les deux dernières étapes prennent place dans l'acinus où on trouve la surface d'échange avec le sang (depuis la 15e génération jusqu'au fond de l'arbre bronchique, 23e génération). Pour que l'acinus soit efficace, l'oxygène doit atteindre les toutes dernières bronches, même si une part de celui–ci est absorbé le long de son parcours de l'arbre. C'est ce que nous avons appelé les effets de screening ou d'écrantage. Cet effet est lié

à la diffusivité de l'oxygène dans l'air, à la perméabilité et à la géométrie de l'arbre bronchique. Pour le poumon humain, il y a typiquement deux situations pour lesquelles on observe des effets d'écrantage :

(a) pour une géométrie donnée, les effet d'écrantage apparaissent lorsque la perméabilité est grande, *i.e.* lorsqu'il est facile pour l'oxygène de traverser la membrane formée de l'épithélium pulmonaire ;

(b) pour une perméabilité donnée, les effets d'écrantage apparaissent lorsque l'acinus est assez grand.

Néanmoins, pour une géométrie et une perméabilité données, ce qui est notre cas pour le poumon, les effets d'écrantage peuvent disparaitre si la source de diffusion est déplacée plus profondément dans les voies aériennes. Ceci arrive dans le poumon lorsqu'on se place à l'exercice. C'est ce qu'on veut observer avec notre modèle.

L'étude et la modélisation de l'hétérogénéité des échanges gazeux dans le poumon n'est pas une problématique tout à fait nouvelle, et il existe une bibliographie assez importante allant de la simple description des phénomènes jusqu'à la proposition de modèles traduisant de manière mathématique ces observations.

Les travaux les plus anciens dont nous avons connaissance sont ceux de PAIVA [Pai73]. Dans ce travail, PAIVA construit un modèle de transport pour les gaz dans le poumon. Il voit le domaine des voies aériennes comme un cylindre de révolution dont une coupe selon l'axe aurait une forme de trompette. Ce domaine est calibré à l'aide des données de WEIBEL. Son modèle est fondé sur les hypothèses suivantes :

(a) les données de WEIBEL sont des moyennes, et la longueur des bronches suit une loi normale de variance égale à la moyenne ;

(b) le diamètre des bronches est suffisamment petit pour considérer la diffusion perpendiculaire à l'axe comme instantanée ;

(c) la diffusion est indépendante de la position géographique dans le poumon.

Pour le domaine géométrique, obtenu à l'aide de la première hypothèse, il distingue le volume occupé par les bronches de celui occupé par les alvéoles, et note $S(\ell)$ la section du modèle correspondant aux deux volumes, bronchique et alvéolaire, cumulés tandis qu'il note $s(\ell)$ la section équivalente correspondant uniquement aux bronches, où ℓ désigne la longueur du chemin parcouru dans l'arbre depuis la bouche. Il suppose que les particules de gaz se situant dans le domaine alvéolaire ne se déplacent pas vers le fond de l'arbre. Il obtient ainsi le modèle pour la diffusion suivant :

$$\frac{\partial c}{\partial t} = D\frac{s}{S}\left[\frac{\partial^2 c}{\partial \ell^2} + \frac{1}{s}\frac{\partial c}{\partial \ell}\frac{\mathrm{d}s}{\mathrm{d}\ell}\right].$$

Ce modèle considère qu'une fois qu'une particule est entrée dans la partie alvéolaire, elle n'en sort plus. PAIVA complète se modèle en ajoutant un terme convectif :

$$\frac{\partial c}{\partial t} = D\frac{s}{S}\left[\frac{\partial^2 c}{\partial \ell^2} + \frac{1}{s}\frac{\partial c}{\partial \ell}\frac{\mathrm{d}s}{\mathrm{d}\ell}\right] - \frac{\dot{V}}{S}\frac{\partial c}{\partial \ell},$$

où $\dot{V}$ ($\mathrm{m}^3{\cdot}\mathrm{s}^{-1}$) est le volume d'air traversant toute section par unité de temps, et ce quelle que soit la section (propriété d'incompressibilité). Il résout ensuite numériquement cette équation, en imposant $\dot{V}$ de sorte à mimer un cycle respiratoire. Il observe une hétérogénéité de la concentration en oxygène dans l'arbre bronchique. Ce premier modèle permet de remarquer le rôle important de la géométrie dans la distribution de la quantité d'oxygène dans le poumon, et donc de son absorption. La description 1D du poumon, permet de prendre en compte la géométrie et notamment l'accroissement de la section totale et l'apparition de volume alvéolaire à partir de la 15e génération, sans pour autant gagner trop de complexité mathématique : grâce à l'hypothèse de symétrie de révolution la géométrie 3D est décrite par une équation 1D. Cependant, ce modèle ne permet pas de quantifier l'oxygène absorbé. Également, le scénario de respiration est imposé via le terme $\dot{V}$, ce que nous remplacerons par le modèle mécanique.

Plus particulièrement, les effets de screening ont été étudiés à l'aide de modèles "labyrinthes", notamment par l'équipe de SAPOVAL *et al.* [FFS+05a, FFS05b, SFW02]. Dans leurs travaux, ils introduisent un paramètre Λ

$$\Lambda = \frac{D_{O_2}}{W_{O_2}},$$

où D_{O_2} est le coefficient de diffusion de l'oxygène dans l'air et W_{O_2} la perméabilité de la membrane pour l'oxygène. Λ a la dimension d'une longueur et compare la capacité à atteindre la surface d'échange et la capacité de

traverser celle–ci. Physiquement, si le diamètre du conduit est plus petit que Λ, les effets de screening peuvent être négligés. Ils comparent ensuite ce paramètre à un autre, plus géométrique cette fois :

$$\Lambda' = \frac{S}{L'},$$

où S correspond à la surface d'échange totale et L correspond au diamètre maximal du domaine acinaire. Λ' a également la dimension d'une longueur et correspond au diamètre moyen d'un acinus. Ainsi, ils établissent que la transition entre les deux régimes, avec effets d'écrantage ou sans effets d'écrantage, se situe lorsque $\Lambda = \Lambda'$. Dans [SFW02], les auteurs comparent ces deux paramètres pour différents mammifères et trouvent que, sur ce critère, le poumon humain est conçu de telle sorte à éviter les effets d'écrantage. Partant de ce constat, ils s'intéressent à la géométrie fractale du poumon. Ils cherchent à savoir si l'air frais entrant dans l'acinus atteint l'ensemble de la surface d'échange disponible (pas d'effet d'écrantage), une large part de la surface d'échange (effets d'écrantage partiels) ou uniquement la surface d'échange située à l'entrée (effets d'écrantage omni–présents). Ils font l'hypothèse que même lorsque la perméabilité est grande, une part de la surface d'échange acinaire pourrait ne pas être active, c'est–à–dire non atteinte par de l'air frais ; de plus la proportion de surface passive augmenterait avec la taille de l'acinus. Afin de confirmer leurs hypothèses, ils effectuent dans leurs travaux des simulations numériques sur des domaines possédant une géométrie dite de Hilbert [SFW02], à savoir une structure auto–similaire ressemblant à une labyrinthe. Voir également [FFS$^+$05a, FFS05b] pour une version 3D. Dans [HGC$^+$05], les auteurs réalisent des simulations comparables, utilisant une géométrie de Sierpinski plutôt qu'une géométrie de Hilbert. Dans [HGC$^+$05], Hou *et al.* complètent leur étude numérique par une étude plus analytique ; ils modélisent l'acinus pulmonaire à l'aide de sa dimension fractale d_f, de sa surface S et de la longueur caractéristique Λ'. Ils en déduisent différentes formulation, plus quantitative et précise que dans [SFW02], du flux d'oxygène vers le sang en fonction de la situation de Λ' par rapport à Λ. Ces expressions analytiques sont en accord avec leurs résultats numériques. Toutefois, ces modèles décrivent uniquement la diffusion de l'oxygène dans l'acinus, et ne sont en aucun cas couplés avec ce qu'il peut se passer en amont de l'acinus dans le poumon. En particulier, la concentration en oxygène est décrite par la concentration en O_2 dans le sang, supposée constante et celle à la source de diffusion, supposée constante elle aussi [HGC$^+$05].

4.4 Un modèle 1D

Le modèle précédemment proposé à la Section 4.2 analyse les sacs alvéolaires comme un unique compartiment et ne prend pas en compte l'hétérogénéité de la distribution des alvéoles le long de l'arbre bronchique. Dans cette section, nous proposons une description continue de la concentration en oxygène le long de l'arbre. Ce nouveau modèle utilise une description 1D du poumon et est couplé encore une fois au modèle mécanique décrit à la Section 2.4.1.

4.4.1 Dérivation du modèle

Une description 1D de l'arbre bronchique L'arbre bronchique est considéré symétrique. Il est modélisé par un simple tube de section variable. La section est constante par morceaux, chaque discontinuité correspondant à un noeud de l'arbre. Voir Figure 4.1. Remarquons que cette description 1D du poumon correspond dans l'idée à celle proposée par Paiva [Pai73].

Remarque 11 *En pratique, étant donné que dans ce modèle nous considérons chaque génération comme un tout, l'hypothèse de symétrie n'est pas nécessaire.*

À nouveau, la variation totale de volume est donnée par la modèle mécanique masse–ressort (voir Section 2.4.1) : la position du parenchyme est notée x, sa vitesse $\dot{x}$ se sorte que la variation de volume total soit donnée par $S\dot{x}$ où S est la surface du parenchyme.

Afin de d'obtenir le modèle continu, fixons tout d'abord quelques notations :

- ℓ désigne la position dans l'arbre bronchique, *i.e.* ℓ est la distance à parcourir dans l'arbre pour atteindre la bouche, $\ell \in [0, L]$

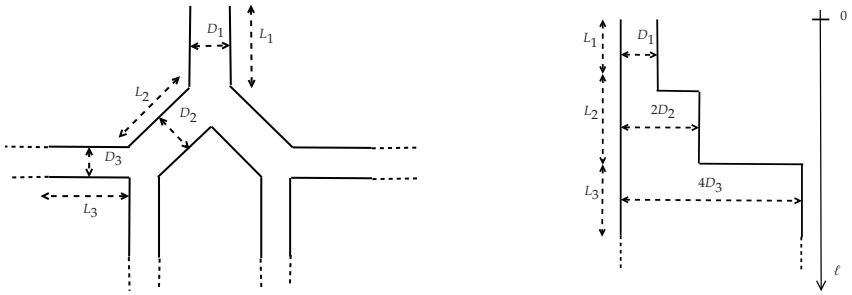


FIG. 4.1: Modèle 1D pour l'arbre bronchique

- $(\ell, t) \mapsto u^{\mathrm{f}}(\ell, t)$ est la vitesse du fluide.

Pour décrire le comportement des bronches, nous noterons :

- $(\ell, t) \mapsto s_b(\ell, t)$ la section totale ;
- $(\ell, t) \mapsto c_b(\ell, t)$ la concentration en oxygène dans la bronche.

L'air dans les alvéoles est décrit par :

- $(\ell, t) \mapsto s_a(\ell, t)$ la section alvéolaire "équivalente", qui est à comprendre comme une densité de volume alvéolaire : $s_a(\ell, t)\,\mathrm{d}\ell$ est le volume alvéolaire compris entre ℓ et $\ell + \mathrm{d}\ell$;
- $(\ell, t) \mapsto c_b(\ell, t)$ la concentration en oxygène dans l'alvéole.

La bronche est connectée aux alvéoles et les alvéoles sont séparés des vaisseaux sanguins par une membrane poreuse. Nous schématisons le modèle à la Figure 4.2. Selon nos notations, $c_a s_a\,\mathrm{d}\ell$ (resp. $c_b s_b\,\mathrm{d}\ell$) est le volume d'oxygène compris entre ℓ et $\ell + \mathrm{d}\ell$ dans les alvéoles (resp. dans les bronches).

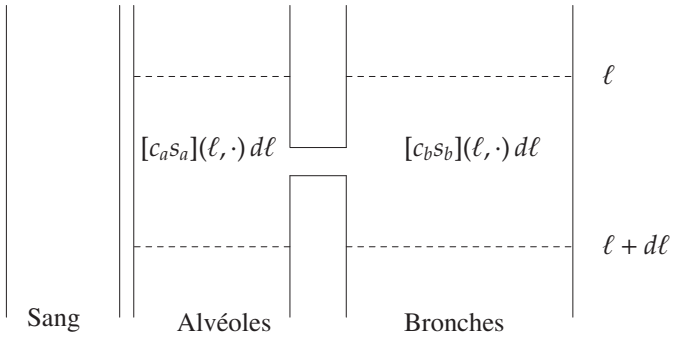


FIG. 4.2: Volumes d'oxygène dans les bronches et les alvéoles.

Le système complet est décrit par le système suivant de trois équations aux dérivées partielles :

$$\partial_t s_b + \partial_\ell(s_b u^{\mathrm{f}}) = -\partial_t s_a, \tag{4.2a}$$

$$\partial_t(s_b c_b) + \partial_\ell(s_b c_b u^{\mathrm{f}}) + \partial_\ell(-D_{O_2} s_b \partial_\ell c_b) = -\partial_t s_a[\dot{x} \geq 0\,?\,c_b, c_a] - W(\ell)\,(c_b - c_a), \tag{4.2b}$$

$$\partial_t(s_a c_a) = -b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}) + \partial_t s_a[\dot{x} \geq 0\,?\,c_b, c_a] + W(\ell)\,(c_b - c_a). \tag{4.2c}$$

L'Équation (4.2a) traduit la conservation du volume d'air sous une hypothèse d'incompressibilité ; $\partial_t s_b$ mesure la variation du volume bronchique et $\partial_t s_a$ celui du volume alvéolaire. La seconde équation, Équation (4.2b), correspond au bilan de masse pour l'oxygène écrit pour les bronches tandis que la troisième, Équation (4.2c), correspond au bilan de masse écrit cette fois pour la alvéoles. Dans le membre de gauche de (4.2b), nous reconnaissons les termes de transport à la vitesse du fluide et de diffusion de l'oxygène.

Le terme commun aux équations (4.2b) et (4.2c) ;

$$\partial_t s_a[\dot{x} \geq 0\,?\,c_b, c_a] + W(\ell)\,(c_b - c_a),$$

correspond aux termes d'échange entre bronche et alvéole ; le terme $\partial_t s_a[\dot{x} \geq 0 \,?\, c_b, c_a]$ est un terme convectif correspondant au volume d'oxygène entrant dans l'alvéole dû à la dépression causée par la déformation mécanique de l'alvéole (à l'inspiration) ou sortant de l'alvéole dû à la surpression causée cette fois par la compression de l'alvéole (à l'expiration). Le terme $W(\ell)\,(c_b - c_a)$ est un terme de diffusion dans la direction 'orthogonale' entre l'alvéole et la bronche.

Le terme de perte dans la troisième équation, Équation (4.2c),

$$-b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}),$$

correspond à la quantité d'oxygène absorbée qui passe dans le sang avec b la perméabilité, $\rho_{V/P}$ le ratio ventilation/perfusion, ℓ_a la densité de surface alvéolaire et $(P_0 c_a - P_{sang})$ la différence de pression partielles en oxygène entre l'alvéole et le sang.

Dans la suite de ce paragraphe, nous détaillons une par une l'obtention de ces trois équations à partir du modèle.

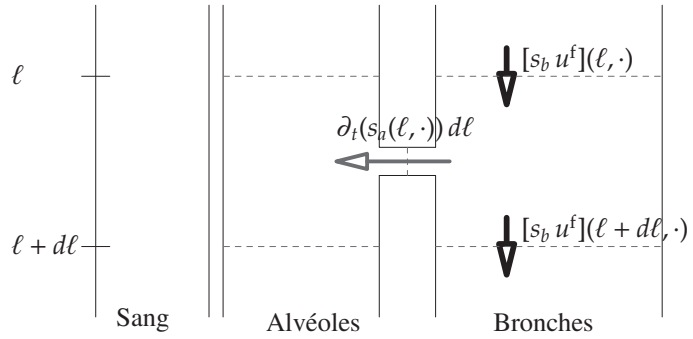


FIG. 4.3: Flux d'air (à l'inspiration)

Conservation du volume d'air La variation totale du volume d'air bronchique compris entre ℓ et $\ell + \mathrm{d}\ell$, $\partial_t(s_b(\ell, \cdot)\,\mathrm{d}\ell)$ est la somme de deux termes :

- la différence entre le flux advectif rentrant et le flux advectif sortant :

$$[s_b u^{\mathrm{f}}](\ell, \cdot) - [s_b u^{\mathrm{f}}](\ell + \mathrm{d}\ell, \cdot),$$

- et la perte de l'air qui s'engouffre dans les alvéoles dû à la déformation de ceux–ci :

$$-\partial_t(s_a(\ell, \cdot)\,\mathrm{d}\ell).$$

Finalement nous obtenons l'Équation (4.2a) :

$$\partial_t s_b c_b + \partial_\ell(s_b u^{\mathrm{f}}) = -\partial_t s_a.$$

Afin de fermer le système, nous ajoutons les conditions aux bords suivantes :

$$\begin{array}{rcl} s_b(0,t)u^{\mathrm{f}}(0,t) & = & S\dot{x}(t), \\ s_b(L,t)u^{\mathrm{f}}(L,t) & = & 0, \end{array} \tag{4.3}$$

où $\dot{x}$ est la vitesse du ressort, donnée par le modèle mécanique.

La variation du volume bronchique $\partial_t s_b$ et la variation du volume alvéolaire $\partial_t s_a$ sont données par des lois pré–établies en fonction de la géométrie qui, à partir de la variation de volume totale donnée par la mécanique, répartissent le volume dans les bronches et les alvéoles. Cette répartition prend en compte le paramètre de rigidité bronchique θ. Pour plus de détails, voir Section 4.7.

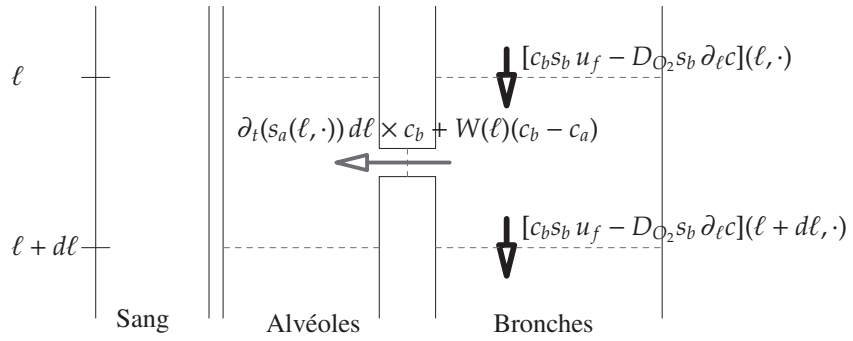


FIG. 4.4: Flux d'oxygène dans la bronche (à l'inspiration)

Bilan de masse pour l'O_2 dans les bronches La variation totale du volume d'oxygène dans la bronche situé entre ℓ et $\ell + \mathrm{d}\ell$, $\partial_t(c_b(\ell, \cdot)s_b(\ell, \cdot)\,\mathrm{d}\ell)$ est la somme de quatre termes :

- le flux advectif :
$$[c_b s_b u^{\mathrm{f}}](\ell, \cdot) - [c_b s_b u^{\mathrm{f}}](\ell + \mathrm{d}\ell, \cdot);$$
- le flux diffusif :
$$[-D_{O_2} s_b \partial_\ell c_b](\ell, \cdot) - [-D_{O_2} s_b \partial_\ell c_b](\ell + \mathrm{d}\ell, \cdot),$$
où D_{O_2} est le coefficient de diffusion de l'oxygène dans l'air ;
- l'oxygène qui entre dans l'alvéole par convection, c'est–à–dire par dépression ou surpression dû à la déformation des alvéoles :
$$\partial_t(s_a(\ell, \cdot)\,\mathrm{d}\ell) \begin{cases} c_b & \text{si } \dot{x} \geq 0 \text{ (inspiration)} \\ c_a & \text{si } \dot{x} < 0 \text{ (expiration)} \end{cases}$$
- et enfin l'oxygène qui entre dans l'alvéole par diffusion :
$$W(\ell)(c_a - c_b).$$
$W(\ell)$ est choisi proportionnel à la densité de surface alvéolaire le long de l'arbre bronchique. Son calcul est détaillé à la Section 4.7.

Finalement, nous obtenons l'Équation (4.2b) :

$$\partial_t(s_b c_b) + \partial_\ell(s_b c_b u^{\mathrm{f}}) + \partial_\ell(-D_{O_2} s_b \partial_\ell c_b) = -\partial_t s_a [\dot{x} \geq 0 \,?\, c_b, c_a]$$

Afin de fermer le système, nous ajoutons les conditions aux bords suivantes :

$$c_b(0, t) = c_0, \tag{4.4}$$

avec $c_0 = 0.2$, la concentration en oxygène dans l'air frais. Nous n'avons pas besoin d'imposer de condition en $\ell = L$ puisque $u^{\mathrm{f}}(L, t) \equiv 0$.

Bilan de masse pour l'O_2 dans les alvéoles La variation totale du volume d'oxygène situé dans les alvéoles compris ℓ et $\ell + \mathrm{d}\ell$, $\partial_t(c_a(\ell, \cdot)s_a(\ell, \cdot)\,\mathrm{d}\ell)$ est la somme de trois termes :

- l'oxygène qui entre dans les alvéoles par convection depuis la bronche, par simple déformation de l'alvéole ce qui crée une dépression :
$$\partial_t(s_a(\ell, \cdot)\,\mathrm{d}\ell)[\dot{x} \geq 0 \,?\, c_b, c_a];$$
- l'oxygène qui entre dans l'alvéole par diffusion :
$$W(\ell)(c_a - c_b)$$

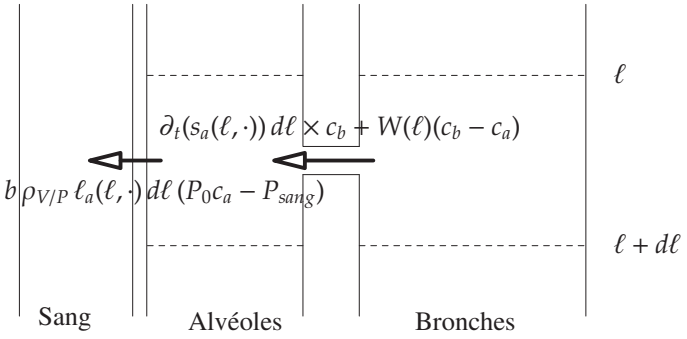


FIG. 4.5: Flux d'oxygène dans les alvéoles(à l'inspiration)

- et l'oxygène qui diffuse à travers la membrane vers le sang :

$$b\,\rho_{V/P}\,\ell_a\,(c_a P_0 - P_{sang}),$$

où

(a) $b\,\ell_a(\ell)$ correspond à la capacité de diffusion de la membrane. $\ell_a(\ell)$ est la densité de surface alvéolaire au repos, déterminée par les valeurs données à la Table 5.1. En d'autres termes, $\ell_a(\ell)\ \mathrm{d}\ell$ est la surface alvéolaire comprise entre ℓ et $\ell + \mathrm{d}\ell$; b est la preméabilité considérée ici comme une constante à caler.

(b) $\rho_{V/P}$ correspond au ratio ventilation/perfusion, il modélise le ratio entre la fréquence respiratoire et la fréquence cardiaque. (Voir Table 5.5 pour les données chiffrées).

(c) P_0 est la pression atmosphérique ($P_0 = 760\,\mathrm{mmHg}$) et P_{sang} la pression partielle en oxygène dans le sang pulmonaire veineux, c'est–à–dire à son entrée dans le système vasculaire pulmonaire. Dans la littérature, nous trouvons $P_{sang} = 40\,\mathrm{mmHg}$.

Finalement, nous obtenons l'Équation (4.2c) :

$$\partial_t(s_a c_a) = -b\,\rho_{V/P}\,\ell_a\,Pe(P_0 c_a - P_{sang}) + \partial_t s_a[\dot{x} \geq 0\ ?\ c_b, c_a] + W(\ell)\,(c_b - c_a).$$

Le système d'équations obtenu est constitué d'équations aux dérivées partielles, que nous allons maintenant discrétiser afin de simuler notre modèle.

4.5 Simulations numériques

Finalement, notre modèle complet, mécanique couplé au modèle 1D pour l'absorption de l'oxygène, est décrit par les équations suivantes :

$$\dot{x} = u, \tag{4.5a}$$

$$\dot{u} = \frac{1}{m}\left(f_{ext} - \left(\frac{R_0 S^2}{1 + \theta S x / V_B^0} + \mu\right)u - kx\right), \tag{4.5b}$$

$$\partial_t s_b + \partial_\ell(s_b u^{\mathrm{f}}) = -\partial_t s_a, \tag{4.5c}$$

$$\partial_t(s_b c_b) + \partial_\ell(s_b c_b u^{\mathrm{f}}) + \partial_\ell(-D_{O_2} s_b \partial_\ell c_b) = -\partial_t s_a[\dot{x} \geq 0\ ?\ c_b, c_a] - W(\ell)\,(c_b - c_a), \tag{4.5d}$$

$$\partial_t(s_a c_a) = -b\,\rho_{V/P}\,\ell_a\,(P_0 c_a - P_{sang}) + \partial_t s_a[\dot{x} \geq 0\ ?\ c_b, c_a] + W(\ell)\,(c_b - c_a), \tag{4.5e}$$

que nous complétons par les conditions aux bords suivantes :

$$\begin{cases} s_b(0,t)u^{\mathrm{f}}(0,t) = S\dot{x}(t), \\ s_b(L,t)u^{\mathrm{f}}(L,t) = 0, \\ c_b(0,t) = c_0, \end{cases} \tag{4.6}$$

et les conditions initiales :

$$\begin{cases} s_b(\ell,0) = s_b^0, \\ c_b(\ell,0) = c_b^0, \quad c_a(\ell,0) = c_a^0, \\ x(0) = x_0, \quad u(0) = u_0. \end{cases} \tag{4.7}$$

Remarque 12 (Couplage faible) *Nous voulons attirer l'attention du lecteur sur le fait que le système* (4.5) *est faiblement couplé : les deux premières équations, les équations mécaniques, déterminent indépendamment du reste de déterminer les valeurs des variables mécaniques x et $\dot{x}$ permettant d'obtenir directement les valeurs de $\partial_t s_b$ et $\partial_t s_a$ données par des lois géométriques ne prenant en argument que la variation de volume soit $S\dot{x}$ (voir Section 4.7). Ainsi, connaisant ces valeurs, la troisième équation permet de déterminer la vitesse du fluide u^{f}. Enfin, connaissant u^{f} les deux dernières équations permettent de déterminer les concentrations c_a et c_b.*

Pour la résolution numérique du système (4.5), nous utilisons une discrétisation par volumes finis pour l'équation (4.5c) en prenant garde de préserver la conservativité. Pour l'Équation (4.5d), nous utilisons le flux upwind pour la partie convective, tandis qu'une discrétisation par différences finies classiques est utilisée pour la partie diffusion. Afin de s'affranchir des contraintes liées à la condition CFL, l'implémentation est entièrement implicite, nécessitant la résolution à chaque pas de temps d'un système linéaire. Cependant, les matrices étant creuses, nous utilisons les routines particulières dédiées au traitement de ce genre de matrices déjà codées dans `Scilab`.

Calcul de ℓ_a Afin de quantifier à quel point les alvéoles sont ouvertes afin de déterminer les surface efficace nous calculons un nombre de Péclet local. Rappelons que le nombre de Péclet mesure quel est le régime : convectif ou diffusif. Il est définit comme suit :

$$Pe := \frac{LU}{D_{O_2}},$$

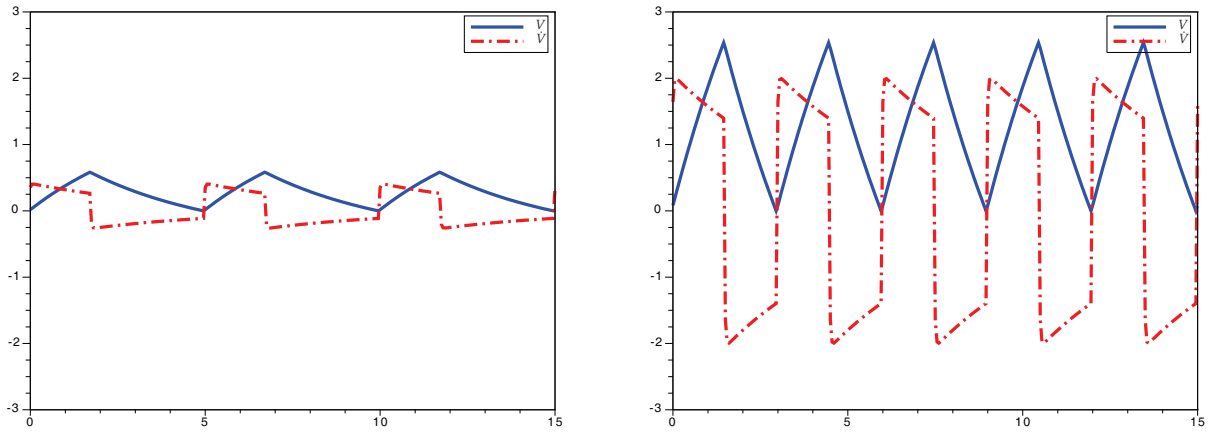
où L correspond à une longueur caractéristique, typiquement la longueur d'une bronche, U à la vitesse caractéristique du fluide et D_{O_2} est le coefficient de diffusion de l'oxygène dans l'air. Pour chaque génération, nous calculons le nombre de Péclet local. Si $Pe > 1$, le régime est essentiellement convectif, il n'y a pas d'écrantage et nous pouvons considérer que toutes les capillaires sont ouverts ; c'est ce qui se passe à l'exercice. Au contraire, si $Pe < 1$ le régime est diffusif et nous devons supposer que la surface d'échange n'est pas disponible dans son ensemble. L'idée de brider la surface disponible par le nombre de Péclet apparaît également dans les travaux de FILOCHE *et al.* [FMAJS08].

4.5.1 Résultats numériques

Tout d'abord, nous donnons à la Figure 4.6, l'évolution du volume bronchique et du flux d'air à la bouche en fonction du temps pour les deux régimes : repos et effort. La description de la force extérieure utilisée dans ces deux cas est décrite à la Section 5.6, page 137.

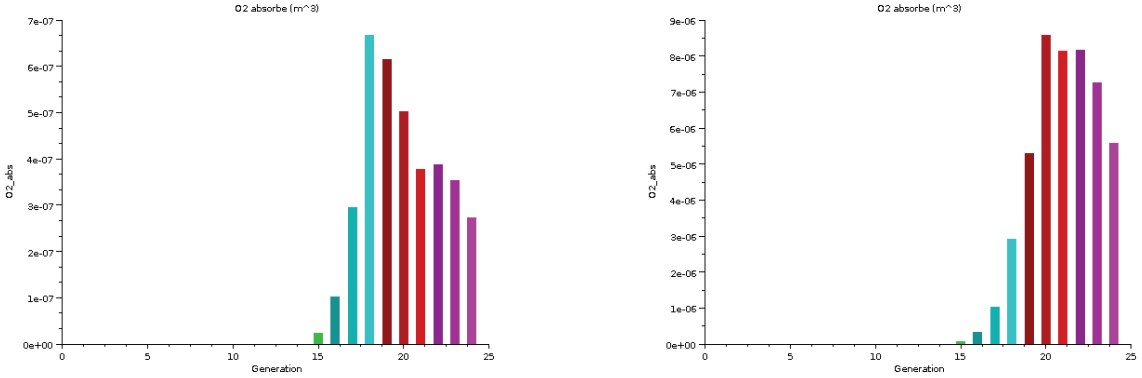
Nous présentons à la Figure 4.7 des histogrammes représentant le flux moyen d'oxygène traversant la membrane par génération, au repos et à l'effort.

Nous notons un décallage vers les bronches les plus basses lors du passage d'un régime de repos à celui de l'exercice, mettant en évidence les effets d'écrantage. Nous donnons également à la Table 4.3, la quantité totale d'oxygène absorbée moyenne par unité de temps. Pour comparaison, nous indiquons les valeurs que l'on trouve dans littérature de physiologie. Nous pouvons consater que les ordres de grandeurs sont bons et nous retrouvons le facteur 3 par la mécanique et la géométrie, ce qui justifie les valeurs de capacité de diffusion données par GUYTON et HALL [GH06]. Nous traçons enfin la concentration alvéolaire moyenne dans chaque génération au cours du temps, voir Figure 4.8 ; notons que même à l'effort, avec ce modèle, la concentration reste élevée même au fond



(a) Repos (b) Effort

FIG. 4.6: Volume pulmonaire et flux d'air à la bouche en fonction du temps.



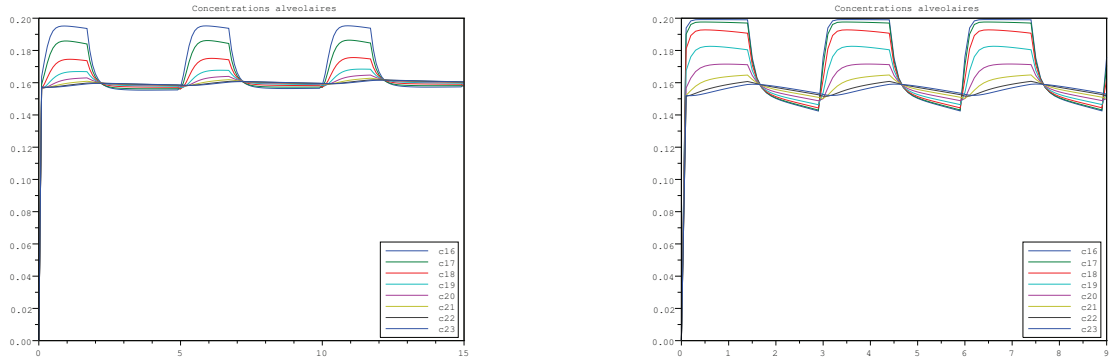
(a) Repos (b) Effort

FIG. 4.7: Flux moyens d'oxygène par génération.

de l'arbre où en pratique il ne devrait pas y avoir renouvellement d'air. C'est un résulat qui d'après nous, ne correspond pas à la réalité et qui joue en défaveur de notre modèle. Ce phénomène est sans doute lié à l'utilisation du nombre de Péclet pour le calcul de la surface efficace : en effet, ce choix n'empêche pas l'oxygène d'arriver aux alvéoles les plus profonds, mais freine juste la diffusion de celui–ci à travers la membrane.

4.5.2 Discussion

L'utilisation du nombre de Péclet dans le calcul de la surface efficace n'est d'après nous pas satisfaisante ; en effet, l'effet du nombre de Péclet devrait déjà être décrit par la prise en compte de la géométrie. Le nombre de Péclet est conséquence de la diminution de la vitesse le long de l'arbre conséquence immédiate de la géométrie en arbre. Cependant, pour l'instant nous ne pouvons nous affranchir de cette description pour trouver les résultats physiologiques que nous nous sommes imposés pour valider notre modèle (voir page 121).



(a) Repos (b) Effort

FIG. 4.8: Concentrations alvéolaires moyennes en oxygène par génération.

	GUYTON & HALL [GH06]	Modèle mathématique 1D
au repos	$4 \cdot 10^{-6}$ m^3 s^{-1}	$3.6 \cdot 10^{-6}$ m^3 s^{-1}
à l'exercice	$35 \cdot 10^{-6}$ m^3 s^{-1}	$47 \cdot 10^{-6}$ m^3 s^{-1}

TAB. 4.3: Flux moyens d'oxygène vers le sang.

Contraintes numériques

Notre arbre est représenté par un domaine 1D où finalement, l'emplacement des noeuds de l'arbre n'a plus d'importance. Toutefois, pour le post–traitement et l'exploitation des résultats, nous réintégrons les quantités obtenues sur chaque bronche (pour obtenir les figures 4.7 par exemple). Afin de garder une précision suffisante, nous devons maintenir un nombre minimum de mailles sur chaque bronche. Ceci couplé avec la condition CFL, nous contraint à l'utilisation d'un schéma implicite pour garder des temps de calculs raisonnables.

De plus, nous sommes confrontés à des problèmes de diffusion numérique : pour le flux upwind, le coefficient de diffusion numérique est de l'ordre $\Delta\ell - u^f\Delta t$, coefficient qui doit être comparé au coefficient de diffusion de l'oxygène dans l'air $D_{O_2} = 0.2 \cdot 10^{-4}$m^2·s^{-1}. Par un petit calcul, nous obtenons rapidement que pour une grille de 1000 mailles, la diffusion numérique est prépondérante. Étant donnée l'importance de la diffusion dans les phénomènes que l'on veut étudier, la diffusion numérique devient vite un problème.

C'est dans cette direction que s'oriente à présent nos travaux ; les pistes envisagées sont :

- ⋆ le raffinement de maillage ;
- ⋆ l'utilisation de schéma d'ordre élevé, tels Lax–Wendroff.

Il reste encore un travail numérique à réaliser sur ce modèle.

4.6 Conclusions et perspectives

Nous avons montré à l'aide d'un modèle 1D l'influence de la géométrie du poumon sur les échanges gazeux, et mis en évidence des effets d'écrantage. Cependant, comme mentionné ci–dessus, ce modèle est critiquable : en effet, l'utilisation du nombre de Péclet pour le calcul de la surface efficace n'est pas satisfaisante car d'après nous, elle réintroduit dans le modèle l'influence de la géométrie de façon assez artificielle. Nous pensons à présent que les résultats numériques obtenus sans ajouter le bridage de la surface par le nombre de Péclet sont entachés d'erreur due à notre schéma numérique. Des travaux allant dans ce sens sont en cours. Il est important de trouver un schéma numérique fiable et assez rapide pour pouvoir ensuite l'utiliser pour une étude appronfondie des paramètres.

En effet, ce modèle correspond à l'étape 0 pour ensuite faire de l'identification de paramètre et étudier l'influence du muscle lisse sur le système. À plus long terme, nous pouvons envisager d'optimiser la force musculaire exercée par le diaphragme pour pouvoir avoir une respiration la plus efficace possible. Une réflexion sera à apporter alors, sur le choix du critère d'efficacité, sans doute différent au repos et à l'effort.

Remerciements

Merci beaucoup à Sébastien d'avoir relu si minutieusement et attentivement ce chapitre.

4.7 Appendix : Détails pour le calcul

4.7.1 Calcul de $\partial_t s_b$ et $\partial_t s_a$

Comme mentionné précédemment, les équations mécaniques (4.5a)–(4.5b), nous donnent la variation totale de volume notée $\dot{V} = S\dot{x}$; ce volume est partagé entre les alvéoles et les bronches selon la loi suivante : $\dot{V}_a(t) = (1-\theta)\dot{V}(t)$ pour les alvéoles et $\dot{V}_b(t) = \theta\dot{V}(t)$ pour les bronches.

- Nous distribuons la variation de volume bronchique $\dot{V}_b$ dans les différentes bronches proportionnellement à la distribution du volume bronchique au repos le long de l'arbre. À cette fin, nous construisons la fonction de distribution $\ell \mapsto \alpha(\ell)$, $\alpha(0) = 0$, $\alpha(L) = 1$ telle que pour chaque bronche :
$$\int_{\ell_{i-1}}^{\ell_i} \frac{\partial \alpha}{\partial \ell}\,\mathrm{d}\ell = \frac{V_b(i)}{\sum_{k=0}^{23} V_b(k)},$$
où $V_b(i)$ est le volume bronchique total situé à la génération i : $V_b(i) = \pi L_i (D_i/2)^2$.
- De la même manière, nous définissons la fonction de distribution $\ell \mapsto \beta(\ell)$, $\beta(0) = 0$, $\beta(L) = 1$ telle que pour chaque génération :
$$\int_{\ell_{i-1}}^{\ell_i} \frac{\partial \beta}{\partial \ell}\,\mathrm{d}\ell = \frac{V_a(i)}{\sum_{k=0}^{23} V_a(k)},$$
où $V_a(i)$ est le volume alvéolaire totale situé à la génération i .

Remarque 13 *En pratique, n'ayant pas d'indications chiffrée sur les volumes alvéolaires mais plutôt sur les volumes alvéolaires on pose $V_a(i) = S_a(i)^{3/2}$, où $S_a(i)$ est la surface alvéolaire par génération, donnée par* WEIBEL *[Wei84], voir Table 5.1, page 135.*

Ainsi, on pose :
$$\partial_t s_b = \dot{V}_b \frac{\mathrm{d}\alpha}{\mathrm{d}\ell} = \theta S\dot{x}\frac{\mathrm{d}\alpha}{\mathrm{d}\ell} \quad \text{et} \quad \partial_t s_a = \dot{V}_a \frac{\mathrm{d}\beta}{\mathrm{d}\ell} = (1-\theta)S\dot{x}\frac{\mathrm{d}\beta}{\mathrm{d}\ell}.$$

4.7.2 Calcul de ℓ_a

La fonction $\ell \mapsto \ell_a$ représente la densité de surface alvéolaire bridée par le nombre de Péclet. Nous construisons tout d'abord une fonction représentant la surface totale disponible à l'aide des données de Weibel ; cette fonction, notée ℓ_a^0 est construite de façon à être continue, monotone et vérifie
$$\int_{\ell_{i-1}}^{\ell_i} \ell_a^0\,\mathrm{d}\ell = S_a(i),$$
où le segment $[\ell_{i-1}, \ell_i]$ est la partie de l'arbre correspondant à la i^e génération et $S_a(i)$ est la surface alvéolaire totale située à la génération i (donnée par WEIBEL, voir voir Table 5.1, page 135). Nous calculons alors le nombre de Péclet local, fonction régularisée de la fonction qui vaut :
$$Pe_i(t) = \frac{L_i U_i(t)}{D_{O_2}},$$
où L_i est la longueur des bronches de la i^e génération, et $U_i(t)$ la vitesse moyenne dans cette génération au temps t. Nous tronquons également cette fonction à 1. Nous obtenons ainsi $Pe(\ell, t) \in [0, 1]$ et nous posons $\ell_a = Pe\,\ell_a^0$.

4.7.3 Calcul de $W(\ell)$

Nous supposons que le terme $W(\ell)$ est proportionnel à la distribution de la surface alvéolaire le long de l'arbre bronchique. Numériquement, il est défini ainsi

$$W(\ell) := W\frac{\partial \gamma}{\partial \ell},$$

où γ est la distribution de surface alvéolaire :

$$\int_{\ell_{i-1}}^{\ell_i} \frac{\partial \gamma}{\partial \ell}\, \mathrm{d}\ell = \frac{S_a(i)}{\sum_{k=0}^{23} S_a(k)}.$$

Chapitre 5

Annexe : Données numériques

5.1 Données morphométriques de WEIBEL pour le poumon

Génération	Longueur (cm)	Diamètre (mm)	Surface alvéolaire (m^2)
0	9.40	18	0
1	4.05	12.2	0
2	1.08	8.3	0
3	0.918	5.6	0
4	0.780	4.5	0
5	0.66	3.5	0
6	0.56	2.8	0
7	0.48	2.3	0
8	0.41	1.86	0
9	0.35	1.54	0
10	0.29	1.30	0
11	0.25	1.09	0
12	0.21	0.95	0
13	0.18	0.82	0
14	0.15	0.74	0
15	0.14	0.50	0.21
16	0.133	0.49	0.69
17	0.12	0.40	2.01
18	0.093	0.38	3.87
19	0.083	0.36	6.57
20	0.070	0.34	10.47
21	0.070	0.31	19.83
22	0.070	0.29	36.12
23	0.070	0.25	81.6

TAB. 5.1: Données de Weibel. Pour les premières générations on donne la moyenne arithmétique sur toutes les bronches d'une même génération.

Remarquons que pour ces données, la distribution de la surface alvéolaire est estimée en supposant que la fraction de la surface alvéolaire totale sur une génération est proportionnel à la fraction de la surface bronchique totale sur cette même génération. Cependant, certains ajustements sont faits sur les trois premières générations, de 15 à 17, afin de prendre en considération le fait que ces générations ne sont pas entièrement recouvertes d'alvéoles (voir Figure 2.2, page 74) [THB08, WSF05].

5.2 Paramètres pour le fluide

5.2.1 Valeurs 3D

viscosité dynamique	$\eta = 1.98 \times 10^{-5}\,\mathrm{Pa}\cdot\mathrm{s}$
densité	$\rho = 1.11\,\mathrm{kg}\cdot\mathrm{m}^{-3}$
viscosité cinématique	$\nu = \eta/\rho = 1.8 \times 10^{-5}\,\mathrm{m}^2\cdot\mathrm{s}^{-1}$

TAB. 5.2: Caractéristiques du fluide réelles, utilisées pour le calcul en dimension $d = 3$.

5.2.2 Valeurs 2D

Pour les calculs en dimension $d = 2$, en ce qui concerne la viscosité du fluide η et la densité du fluide ρ, on ne choisit pas les valeurs physiologiques, mais des valeurs telles que la résistance globale de notre arbre 2D soit égale à la résistance globale de l'arbre réel 3D, et telles que le nombre de Reynolds 2D (voir (2.3), page 81) soit égal au nombre de Reynolds dans le cas 3D (voir [SA07]). En conséquence, les constantes sont

viscosité dynamique	$\eta = 4 \times 10^{-3}\,\mathrm{Pa}\cdot\mathrm{s}$
densité	$\rho = 50\,\mathrm{kg}\cdot\mathrm{m}^{-3}$
viscosité cinématique	$\nu = \eta/\rho = 8 \times 10^{-5}\,\mathrm{m}^2\cdot\mathrm{s}^{-1}$

TAB. 5.3: Paramètres fluides utilisés pour calcul en dimension $d = 2$.

5.3 Paramètres pour le système masse-ressort

Afin de reproduire des courbes expérimentales, nous avons utilisé des données physiologiques, choisies proches de celles que l'on trouve dans [BT06], [MMSS08]

- $m = 0.3\,\mathrm{kg}$, la masse totale du poumon ;
- $S = 0.011\,\mathrm{m}^2$, surface de la face mobile de la boîte (surface du diaphragme) ;
- $E = 3.32 \times 10^5\,\mathrm{N}\cdot\mathrm{m}^{-5}$, l'élastance du poumon ;
- $k_0 = E \cdot S^2 = 40.172\,\mathrm{N}\cdot\mathrm{m}^{-1}$, la constante de raideur du ressort ;
- $R = 3{\times}10^5\,\mathrm{Pa}\cdot\mathrm{s}\cdot\mathrm{m}^{-3}$, en dimension $d = 2$ et $R = 6.59{\times}10^5\,\mathrm{Pa}\cdot\mathrm{s}\cdot\mathrm{m}^{-3}$, en dimension $d = 3$, la résistance de l'arbre bronchique ;
- $\mu = 0.25 \cdot R \cdot S^2$, la résistance des tissus (20% de la résistance totale).

5.4 Modèle non–linéaire : constante de raideur

On rappelle que la constante de raideur est choisie vérifiant :

$$k(x) = k_0 + \begin{cases} (f_{min}/x_{min} - k_0)x/x_{min}, & \text{si} \quad x < 0 \\ (f_{max}/x_{max} - k_0)x/x_{max}, & \text{si} \quad x \geq 0 \end{cases}.$$

On fixe :

$$\begin{aligned} f_{min} &= -11\,\mathrm{N}, \\ f_{max} &= 13\,\mathrm{N}. \end{aligned}$$

et

$$\begin{aligned} x_{min} &= -0.25\,\mathrm{m}, \\ x_{max} &= 0.20\,\mathrm{m}. \end{aligned}$$

5.5 Volumes bronchiques

Voici les volumes bronchiques utilisés dans les calculs :

- **Volume alvéolaire au repos** : $V_A^0 = 3 \times 10^{-3}\,\mathrm{m}^3$;
- **Volume bronchique au repos** : $V_B^0 = 5 \times 10^{-4}\,\mathrm{m}^3$.

5.6 Trois scénarii pour la force extérieure

La force extérieure est décrite dans tous les cas par une fonction créneau. Dans le cas de la respiration au repos comme à l'effort c'est une fonction T–périodique. α désigne la fraction du cycle consacrée à l'inspiration.

5.6.1 Au repos

Au repos, on suppose que le sujet n'exerce pas de force à l'expiration, le retour à la position d'équilibre se faisant grâce à la force de retour élastique du système. On choisit dans nos simulations une période $T = 5\,\mathrm{s}$ et $\alpha = 0.35$. La force est donnée par la fonction T–périodique vérifiant :

$$f_{ext}(t) = \begin{cases} 2\,\mathrm{N} & \text{si } t \in [0, \alpha T], \\ 0\,\mathrm{N} & \text{si } t \in [\alpha T, T]. \end{cases}$$

5.6.2 A l'exercice

Au contraire, à l'exercice, la force de retour élastique n'est pas suffisante pour obtenir une expiration rapide et efficace et une force suplémentaire est exercée à l'expiration. De plus le cycle est raccourci, $T = 3\,\mathrm{s}$ et le temps d'expiration devient égal à celui de l'inspiration, $\alpha = 0.5$ [Lag]. La force est alors donnée par la fonction T–périodique vérifiant :

$$f_{ext}(t) = \begin{cases} 9\,\mathrm{N} & \text{si } t \in [0, \alpha T[, \\ -4\,\mathrm{N} & \text{si } t \in [\alpha T, T[. \end{cases}$$

5.6.3 Manoeuvre forcée

Comme expliqué au Chapitre 2 (page 76), au cours d'un test de spirométrie le patient exécute une manoeuvre particulière. Afin de reproduire ces expériences *in silico*, on utilise l'expression de la force extérieure exercée suivante [BGM09]. La manoeuvre dure une vingtaine de seconde et on a

$$f_{ext}(t) = \begin{cases} 4\,\mathrm{N} & \text{si } t \in [0, 0.5[, \\ -1\,\mathrm{N} & \text{si } t \in [0.5, 2[, \\ 2\,\mathrm{N} & \text{si } t \in [2, 2.5[, \\ -1.5\,\mathrm{N} & \text{si } t \in [2.5, 3[, \\ 3\,\mathrm{N} & \text{si } t \in [3, 4[, \\ -1\,\mathrm{N} & \text{si } t \in [4, 5[, \\ 2\,\mathrm{N} & \text{si } t \in [5, 6[, \\ -1\,\mathrm{N} & \text{si } t \in [6, 8[, \\ a\,t^3 + b\,\mathrm{N} & \text{si } t \in [8, 9[, \\ f_{max}\,\mathrm{N} & \text{si } t \in [9, 10.5[, \\ f_{min}\,\mathrm{N} & \text{si } t \in [10.5, 17[, \\ c\,t^3 + d\,\mathrm{N} & \text{si } t \in [17, 18[, \\ 10\,\mathrm{N} & \text{si } t \in [18, 20[, \end{cases}$$

où a, b, c et d sont choisis de façon à ce que f_{ext} soit continue et

$$\begin{aligned} f_{min} &= -11\,\mathrm{N}, \\ f_{max} &= 13\,\mathrm{N}. \end{aligned}$$

Cette fonction est représentée à la Figure 3.4.

5.7 Volumes d'oxygène absobé

Les volumes moyens d'oxygène absorbé sont mesurés exprérimentalement. On résume à la Table 5.4 les valeurs que l'on peut trouver dans la littérature de physiologie.

Régime	repos	effort
$\langle \dot{V}_{O_2} \rangle$	$4 \cdot 10^{-6}\,\mathrm{m}^3 \cdot \mathrm{s}^{-1}$ [Wei84, GH06]	entre $35 \cdot 10^{-6}$ [Wei84] et $54 \cdot 10^{-6}\,\mathrm{m}^3 \cdot \mathrm{s}^{-1}$[SFW02]

TAB. 5.4: Volumes moyen d'oxygène absorbé.

5.8 Ratio ventilation/perfusion

Le ratio ventilation/perfusion mesure le rapport entre le rythme respiratoire et le rythme cardiaque. Nous donnons dans la Table 5.5 les valeurs que l'on peut trouver dans [GH06]

Régime	repos	effort
$\rho_{V/P}$	1/6	1/2

TAB. 5.5: Ratio ventilation/perfusion.

Bibliographie

[ABF84] Douglas N. Arnold, Franco Brezzi, and Michel Fortin.
A stable finite element for the Stokes equations.
Calcolo, 21(4) :337–344 (1985), 1984.

[Bat70] George Keith Batchelor.
An introduction in fluid dynamics.
U.P., Cambridge, 1970.

[Bat09] Jason H.T. Bates.
Lung Mechanics, an inverse modeling approach.
Cambridge University Press, 2009.

[BF91] Franco Brezzi and Michel Fortin.
Mixed and hybrid finite element methods, volume 15 of *Springer Series in Computational Mathematics*.
Springer-Verlag, New York, 1991.

[BGM09] Leonardo Baffico, Céline Grandmont, and Bertrand Maury.
Multiscale modelling of the respiratory tract.
Math. Models Methods Appl. Sci., 2009.

[BT06] Alona Ben-Tal.
Simplified models for gas exchange in the human lungs.
J. Theor. Biol., 238 :474–495, 2006.

[Cho68] Alexandre J. Chorin.
Numerical solution of the Navier–Stokes equations.
Math. Comp., 22 :745–762, 1968.

[Cho69] Alexandre J. Chorin.
On the convergence of discrete approximations to the Navier–Stokes equations.
Math. Comp., 23 :341–353, 1969.

[DGG+09] Anne Devys, Céline Grandmont, Bérénice Grec, Bertrand Maury, and Driss Yakoubi.
Numerical method for the 2d simulation of the respiration.
ESAIM : Proc., 28 :162–181, 2009.

[FFS+05a] Maddalena Felici, Marcel Filoche, Christian Straus, Thomas Similowski, and Bernard Sapoval.
Diffusional screening in real 3D human acini–a theoretical study.
Respir Physiol Neurobiol, 145(2-3) :279–93, 2005.

[FFS05b] Marcel Filoche, Maddalena Felici, and Bernard Sapoval.
Renormalized random walk study of oxygen absorption in the human lung.
Physica A : Statistical Mechanics and its Applications, 357(1) :27, 2005.

[FGNQ02] Luca Formaggia, Jean-Frédéric Gerbeau, Fabio Nobile, and Alfio Quarteroni.
Numerical treatment of defective boundary conditions for the Navier-Stokes equations.
SIAM J. Numer. Anal., 40(1) :376–401 (electronic), 2002.

[FMAJS08] Marcel Filoche, André A. Moreira, José S. Andrade Jr., and Bernard Sapoval.
Quantitative analysis of the oxygen transfer in the human acinus.
Integration In Respiratory Control : From Genes To Systems, 605 :167–172, 2008.

[FMP+05] Catalin Fetita, Simona Mancini, Diane Perchet, Françoise Prêteux, Marc Thiriet, and Laurence Vial.

An image based computational model of oscillatory flow in the proximal part of tracheobronchial trees.
Comput. Meth. Biomech. Biomed. Eng., 8, 2005.

[GH06] Arthur C. Guyton and John E. Hall.
Textbook of Medical Physiology, 11th–edition.
Elsevier Saunder, 2006.

[GMM05] Céline Grandmont, Yvon Maday, and Bertrand Maury.
A multiscale/multimodel approach of the respiration tree.
In *New trends in continuum mechanics*, volume 3 of *Theta Ser. Adv. Math.*, pages 147–157. Theta, Bucharest, 2005.

[GMS08] Céline Grandmont, Bertrand Maury, and Assia Soualah.
Multiscale modelling of the respiratory track : a theoretical framework.
ESAIM Proc., 23 :10–29, 2008.

[HGC+05] Chen Hou, Stefan Gheorghiu, Marc-Olivier Coppens, Virginia H. Huxley, and Peter Pfeifer.
Fractals in biology and medicine.
In *Gas Diffusion through the Fractal Landscape of the Lung : How Deep Does Oxygen Enter the Alveolar System ?*, pages 17–30, 2005.

[HLHOP] Frédéric Hecht, Antoine Le Hyaric, Koji Ohtsuka, and Olivier Pironneau.
Freefem++, finite elements software, `http ://www.freefem.org/ff++/`.

[Lag] Frédéric Lagoutière.
Personal communication.

[LTMH07] C.-L. Lin, M. H. Tawhai, G. McLennanc, and E.A. Hoffman.
Characteristics of the turbulent laryngeal jet and its effect on airflow in the human intra-thoracic airways.
Respiratory Physiology & Neurobiology, 157 :295–309, 2007.

[Mau04] Benjamin Mauroy.
Hydrodynamics in the lungs, relations between flows and geometries.
PhD thesis, ENS Cachan, 2004.

[MFWS04] Benjamin Mauroy, Marcel Filoche, Ewald R. Weibel, and Bernard Sapoval.
An optimal bronchial tree may be dangerous.
Nature, 427 :633–636, February 2004.

[MMSS08] Sébastien Martin, Bertrand Maury, Thomas Similowski, and Christian Straus.
Impact of respiratory mechanics model parameters on gas exchange efficiency.
ESAIM : PROCEEDINGS, 23 :30–47, 2008.

[Olu99] Mette Sofie Olufsen.
Structured tree outflow condition for blood flow in larger systemic arteries.
Am. J. Physiol., 276 :257–H268, 1999.

[Pai73] Manuel Paiva.
Gas transport in the human lung.
J Appl Physiol, 35(3) :401–410, 1973.

[Pir88] Olivier Pironneau.
Méthodes des éléments finis pour les fluides.
Paris ; Milan ; Barcelone ; Mexico : Masson, 1988.

[QRV01] Alfio Quarteroni, Stefania Ragni, and Alessandro Veneziani.
Coupling between lumped and distributed models for blood flow problems.
Comput.Visualization Sci., 4 (2) :111124, 2001.

[Qua93] Luigi Quartapelle.
Numerical solution of the incompressible Navier–Stokes equations.
International Series of Numerical Mathematics, 113, 1993.

[QV03] Alfio Quarteroni and Alessandro Veneziani.
Analysis of a geometrical multiscale model based on the coupling of odes and pdes for blood flow simulations.

Multiscale Model. Simul., 1, No.2 :173–195, 2003.

[RHT96] Rolf Rannacher, John G. Heywood, and Stefan Turek.
Artificial boundaries and flux and pressure conditions for the incompressible navier-stokes equations.
International Journal for Numerical Methods in Fluids, 22 :325–352, 1996.

[SA07] Assia Soualah-Alila.
Modélisation mathématique et numŕique du poumon humain.
PhD thesis, Université Paris–Sud XI, 2007.

[SFW02] Bernard Sapoval, Marcel Filoche, and Ewald R. Weibel.
Smaller is better–but not too small : a physical scale for the design of the mammalian pulmonary acinus.
Proc. Natl. Acad. Sci. U.S.A., 99(16) :10411–6, 2002.

[Tem77] Roger Temam.
Navier–Stokes equations. theory and numerical analysis.
Studies in Mathematics and its Applications, 2, 1977.

[THB08] Akira Tsuda, Frank S. Henry, and James P. Butler.
Gas and aerosol mixing in the acinus.
Respiratory Physiology & Neurobiology, 163(1-3) :139 – 149, 2008.
Respiratory Biomechanics.

[VCFJT06] Irene E. Vignon-Clementel, C. Alberto Figueroa, Kenneth E. Jansen, and Charles A. Taylor.
Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries.
Comput. Methods Appl. Mech. Engrg., 195(29-32) :3776–3796, 2006.

[VV07] Alessandro Veneziani and Christian Vergara.
An approximate method for solving incompressible Navier-Stokes problems with flow rate conditions.
Comput. Methods Appl. Mech. Engrg., 196(9-12) :1685–1700, 2007.

[Wei84] Ewald R. Weibel.
The pathway for oxygen : structure and function in the mammalian respiratory system.
Harvard University Press, 1984.

[Wes79] John B. West.
Respiratory physiology : the essentials.
Williams & Wilkins, Baltimore :, 1979.

[WG62] Ewald R. Weibel and Domingo M. Gomez.
Architecture of the Human Lung : Use of quantitative methods establishes fundamental relations between size and number of lung structures.
Science, 137(3530) :577–585, 1962.

[WSF05] Ewald R. Weibel, Bernard Sapoval, and Marcel Filoche.
Design of peripheral airways for efficient gas exchange.
Respiratory Physiology & Neurobiology, 148(1-2) :3 – 21, 2005.
Structure and Function in the Periphery of the Lung.

Troisième partie

Modélisation d'une cascade protéinique

Chapitre 6

Propagation d'un signal dans la cascade MAPK : rôle de la bistabilité et de l'ultrasensitivité pour un problème mixte

La voie de signalisation MAPK est de nos jours décrite [Rus09] comme un réseau de modules communs à de nombreuses espèces. Sa fonction principale est de transférer un signal de la membrane plasmique (*i.e.* membrane cellulaire) à la membrane nucléaire. A la question « comment » répondent les notions d'ultrasensitivité et de bistabilité, liées à la dynamique non–linéaire du système biochimique, ignorant tout phénomène spatial. En ce qui concerne la propagation du signal à travers le cytoplasme, les choses sont bien moins connues. Dans le travail qui suit, fruit d'une collaboration avec RALF BLOSSEY de l'IRI (Institut de Recherche Interdisciplinaire, Lille) d'une part et THIERRY GOUDON et PAULINE LAFITTE d'autre part, nous proposons à partir des lois cinétiques de Michaelis–Menten un modèle de réaction diffusion pour la cascade MAPK dans les ovocytes de Xénopes. Nous étudions ce modèle en 1D. En nous fondant sur des résultats théoriques existants [AW75], nous nous attachons particulièrement aux conditions nécessaires pour la propagation du signal. Nous montrons que l'existence d'un front dépend sensiblement des conditions initiales et au bord (*i.e.* à la membrane plasmique). Enfin, les conséquences biologiques sont discutées. Ce travail a été soumis pour publication sous le titre *Signal propagation in the MAPK cascade : role of bistability and ultrasensitivity for a mixed problem*.

6.1 Introduction and biological context

In this section we give the biological context that motivated this work. This is also the occasion to recall some basics of biology that can be helpful to understand the framework and the challenges that are behind. Note that this work benefits from a collaboration with Ralf Blossey from IRI (Interdisciplinary Research Institute).

Among the large family of mitogenic signals, the signal chain MAPK plays a crucial role in many parts of the cell life : cell cycle, differentiation, angiogenesis ... In particular, it is one of the targets of some strategies against cancer.

A quite important research both experimental and theoretical, is conducted in this direction ; in Lille, the team of Bodart *et al.* ([RGBB09, Rus09]) works on this topic. It is partially funded by the INCa (Institut National du Cancer).

6.1.1 The cell cycle

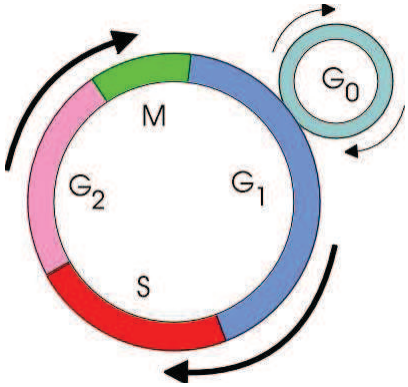


Figure 6.1: The cell cycle

The cell cycle of an eukaryote cell (cell with a nucleus, mamalian cells for instance) is a sequence of steps that leads to the division of the mother cell into two daughter cells. Each step needs the achievement of the previous one. During a cell cycle there are four main steps : the G1 phase, the S phase, the G2 phase and the M phase. Together, G1, S and G2 form the interphase; the M phase is the phase during which the division takes place. During the G1 phase, the cell grows up and get ready for the S phase. The ensuing S phase starts when DNA synthesis begins; when it is complete, all of the chromosomes have been replicated, *i.e.*, each chromosome has two (sister) chromatids. The G2 phase is similar to the G1 phase in the sense that the cell gets ready to divide itself. The M phase is the mitosis : it consists of nuclear division and cytoplasmic division. It is itself divided into four steps : prophase, metaphase, anaphase and telophase. At the end of these steps, the daughter cells can either go into a new cell cycle or enter a quiescent phase, called the G0 phase which is a phase during which basically nothing happens. See Figure 6.1

The molecular mecanisms are essentially the same in all eukaryote cells. The events of the cell divisions are governed by the periodic activations of kinases (*i.e.* proteins) which are part of a very complex network which brings into play a lot of molecules (enzymes, cyclines ...). Research on the biological control of the cell cycle won Tim Hunt, Leland Hartwell and Paul Nurse the Nobel prize in 2001.

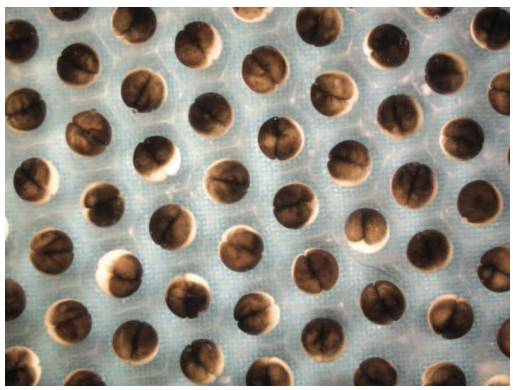
Some kinases have the task to control the switch from one phase to an other. The *Mitogen–Activated Protein Kinases* (MAPK) are one of those, whereby a mitogen is a factor that stimulates mitosis. They have an important role for the transition G1/S, but they have also an impact on the transition G2/M, which does not affect the genetic material. They also have a crucial role in the formation of the mitotic spindle (which is an array of microtubules organizing and separating the chromosomes during mitosis).

6.1.2 *Xenopus* oocyte

The biological context that underlies this work concerns the study of the *Xenopus* oocytes. Among the different models that are available for the biologists, the oocyte of this amphibian is well adapted to the study of the cell cycle : it is easy to synchronise the cells (in G2 phase or in metaphase II during the meiose), it produces a huge quantity of proteins (200 to 400 ng per day) and the size of the cells (1.2 to 1.4 mm of diameter) allows to manipulate them quite easily.



(a) *Xenopus Laevis*



(b) *Xenopus* oocytes

Figure 6.2: *Xenopus laevis* (From `www.euroherp.com` and `www.unc.edu/~fconlon/Daniel/`)

Xenopus laevis is a species of South African aquatic frog (see Figure 6.2). It can grow up to a length of 12 cm with a flattened head and body, but no external ear or tongue. They have an external reproduction mode : they lay

eggs in clusters which are then fertilized.

6.2 Formulation of the problem

In general, a signal (caused *e.g.* by a hormone) arriving at a cell membrane needs to be transferred to the cell interior or its nucleus in order to excite a coherent cellular response [KHK10]. Dedicated signalling chains have evolved to achieve this task. In what follows we are particularly interested in a prominent one, the MAP kinase signalling chain [Kho02]. The acronym MAP stands for mitogen-activated protein, whereby a mitogen is a factor that stimulates mitosis. Upon arrival of a hormone at the cell, generally via a receptor in the cell membrane, a cascade of molecular activation events is launched involving the phosphorylation of downstream effectors. The MAPK cascade is prominent in this respect since it occurs in multiple different organisms and cell types [KHK10]. While the main cascade motif is universal, its embedding into a cell-type and organism-specific signalling network varies. This concerns, *e.g.*, the initiation of the cascade which in somatic cells depends on a well-characterised ensemble of molecules, notably Ras and Raf [Kho02]. Here, we are concentrating on the cascade in the egg of the Southern African clawed frog *Xenopus* laevis, where it is involved in egg maturation. The main initiator of the cascade in *Xenopus* oocytes is the molecule Mos, whose production by translation from a pre-established mRNA is initiated by a hormonal stimulation [FJPK+09]. Figure 1 shows a schematic representation of the cascade and the intervening molecules, following Figure 2 from [FJPK+09].

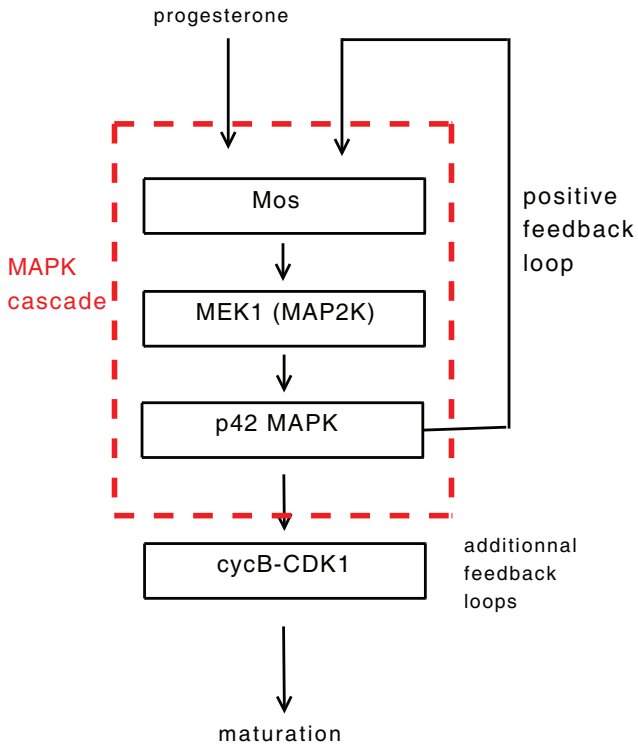


Figure 6.3: Schematic representation of the MAPK cascade in *Xenopus* oocytes. A hormonal signal (progesterone) stimulates the membrane receptors. The signal activates the translation of Mos from a pre-established stock of mRNA. The cascade evolves via the double phosphorylation of MEK1 (MAP2K) and p42-MAPK. A positive feedback loop stimulates Mos activation. The MAPK signal at cascade output is processed further in cell cycle events downstream of the cascade, indicated by Cyclin-B and Cdk1, the maturation- or mitosis-promoting factor, also called MPF.

A particularly noteworthy feature of the MAPK cascade in *Xenopus* oocytes is its feedback control, which has been disentangled in great detail by FERRELL, JR. *et al.*, in both experiment and theory (see [FJPK+09] and

references therein). The most basic feature that we rely on in our work is the positive feedback loop at the cascade exit, p42 MAPK (MAPK), which stimulates the activity of the molecule at cascade entry, Mos (see Figure 6.3).

While a whole body of work exists on the biochemical kinetics of the cascade, far less is known so far about how the input signal traverses the intracellular space. Depending on cell type, at least three mechanisms are known to play a role. These are diffusion in the cytoplasm, the kinetics of the phosphorylation reactions, and, essentially in somatic cells, scaffold proteins [BS09]. Recent reviews on the state of the field are given by [KHK10, Kho06], which give exhaustive lists of references.

A standard approach in the mathematical modeling of the spatio-temporal evolution of the MAPK cascade is via diffusion-reaction equations. For the MAPK casacade, this approach has been most prominently pursued by KHOLODENKO *et al.* [Kho06, KHK10, MTHK06]. In our work, we follow this path, but since we are interested in signal propagation in *Xenopus* oocytes, we can make use of a number of simplifications which allow a more in-depth mathematical analysis. The mathematical simplifications, detailed below, are based on the fact that *Xenopus* oocytes are large cells [FJPK$^+$09] : while at the beginning of the life cycle they are comparable to somatic cells, in their development they pass into a growth phase before maturation is initiated. Cell volumes reach 1 μL with protein concentrations of 25 μg. Since scaffold proteins are not known to play a role in *Xenopus* oocytes, we can consider the propagation to be well-described by a deterministic diffusion-reaction.

Formulation of the model

The mathematical model we develop is based on the usual reaction kinetics of the MAPK cascade in which an upstream kinase activates a downstream one at two residues. For this kinetics, either a mass-action law or a Michaelis-Menten kinetics can be assumed. The kinetic model is described in more detail in Appendix A (Section 6.7). In this way one obtains the usual system of ordinary differential equations for the concentrations of the molecules in the cascade, together with conservation laws for their total concentrations.

In order to arrive at a spatial model we require the following assumptions.

Assumption 1. We assume that molecular transport in the cytoplasm does not involve the action of specialised protein machineries in the cell. This is an acceptable assumption for the *Xenopus* oocytes, but cannot be generalized to somatic cells. As a result of this assumption, the main mechanism of transport is molecular diffusion.

Assumption 2. We assume that the chemical reactions (phosphorylations) in the cascade are always to be considered as fast in the sense that at every instant in time, the cascade molecules are in chemical equilibrium. For a non-compartimentalized cell of high protein content such as *Xenopus* oocytes, this is a reasonable starting assumption, which again might not be easily generalizable to other cell types.

Starting from these assumptions, let us now derive an equation for the concentration of the molecule at the cascade entry, that is Mos. To this end, let $m(x,t)$ be the concentration of molecule Mos. We proceed in two steps :

1. Thanks to Assumption 2 and omitting at first the feedback loop (represented by the red dash line in Figure 6.4), we can use the Michaelis–Menten kinetics to express the concentration of all molecular species as simple functions of the molecule at cascade entry which initialize the signaling chain. Neither MAP2K nor MAPK are therefore considered as diffusing quantities.

2. We take into account the retroaction of z_2 on the concentration of Mos to obtain a spatial model for the concentration of Mos.

The expression of the reaction term depends on the kinetics chosen for the signaling cascade. Typically, the cascade has two phosphorylation states so that for a two-level cascade with molecules y_0 and z_0, two respective phosphorylated states exist, for y_0 (resp. z_0) these are called y_1 and y_2 (resp. z_1 and z_2) (See Figure 6.4). Two particular types of kinetics are linear (the usual law of mass action) and saturated, *i.e.* Michaelis-Menten kinetics.

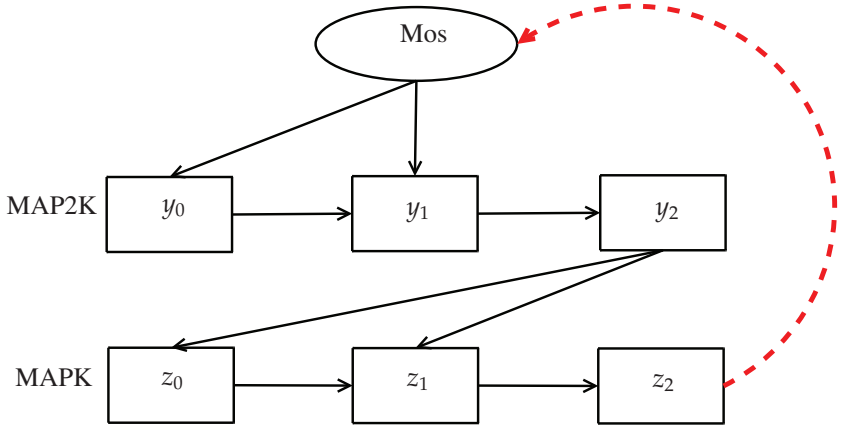


Figure 6.4: The two–level phosphorylation

In the latter case, the equations read following [AFS04] :

$$\dot{y}_0 = \frac{V_6 y_1}{K_6 + y_1} - \frac{V_3 m y_0}{K_3 + y_0} \tag{6.1}$$

$$\dot{y}_1 = -(\dot{y}_0 + \dot{y}_2) \tag{6.2}$$

$$\dot{y}_2 = \frac{V_4 m y_1}{K_4 + y_1} - \frac{V_5 y_2}{K_5 + y_2} \tag{6.3}$$

$$\dot{z}_0 = \frac{V_{10} z_1}{K_{10} + z_1} - \frac{V_7 y_3 z_0}{K_7 + z_0} \tag{6.4}$$

$$\dot{z}_1 = -(\dot{z}_0 + \dot{z}_2) \tag{6.5}$$

$$\dot{z}_2 = \frac{V_8 y_3 z_1}{K_8 + z_1} - \frac{V_9 z_2}{K_9 + z_2} \tag{6.6}$$

where the V_i and K_i are reaction speeds and equilibrium constants, respectively, and where the numbering of the reactions follows the scheme by [Kho00] which is also used in [AFS04]. This scheme simply numbers the reactions sequentially layer by layer, first all phosphorylations, and then all dephosphorylations. Note that if equilibrium constants are much larger than the molecular concentrations, the linear regime is recovered. Equations (6.2) and (6.5) are the result of the matter conservation.

As shown in [RGBB09], for a particular symmetric choice of the parameters, the system can be rendered non-dimensional. For this parameter-symmetric case, the introduction of the functions

$$w_i \equiv \frac{y_i}{1 + y_i}\,, \quad q_i \equiv \frac{z_i}{1 + z_i}\,.$$

allows to rewrite the right–hand side of the equations in a simple polynomial form, from which the fixed-point value conditions are easy to read off. For the variable w_i, they are given by $w_2 = m w_1 = m^2 w_0$.

Exploiting the conservation of molecules y_i as expressed by the condition $\sum_{i=0,1,2} y_i = y_T$, one finds after suitable normalization the condition

$$\frac{w_0}{1 - w_0} + \frac{w_1}{1 - w_1} + \frac{w_2}{1 - w_2} = 1\,.$$

which at the fixed-point reduces to a cubic equation, *e.g.* for w_0,

$$4m^3 w_0^3 - 3m(1 + m + m^2) w_0^2 + 2(1 + m + m^2) w_0 - 1 = 0\,. \tag{6.7}$$

which can be solved exactly. The case of linear kinetics can, however, also be recovered by ignoring the highest order terms of this equation. One then finds

$$w_0 \approx \frac{1}{2}\frac{1}{1+m+m^2}\,. \tag{6.8}$$

Although the full final expression for w_0 from the cubic equation does have a more involved form, the degree of the polynomial remains 2 as for the linear kinetics. From the fixed-point relation, we have

$$w_2 \approx \frac{1}{2}\frac{m^2}{1+m+m^2}\,. \tag{6.9}$$

The same calculation can be repeated for the variable q_i. One obtains an equation of the same form as (6.63) in the variable q_1, only with m replaced by

$$s = v\frac{w_2}{1-w_2}$$

where $v \equiv V_7/V_{10}$. The same argument can then be repeated leading to

$$q_2 \approx \frac{1}{2}\frac{s^2}{1+s+s^2}\,. \tag{6.10}$$

Putting the equations (6.65) and (6.67) together then leads already for the approximate linear kinetics to an involved expression, however with a polynomial of maximal degree four in the variable m in both denominator and numerator. This remains the case also after the required transforms back from w_i and q_i to the original variables y_i and z_i. While the full expression obtained from the two cubic equations can still be written down analytically, it turns out to be yet more involved and less illuminating. At the end, we get

$$z_2 \approx \frac{m^4}{1+m^2+m^4}$$

Knowing the expression for the concentration of z_2, we can propose a model for the transport of Mos. We have

$$\partial_t m(x,t) - D\Delta m(x,t) = R[m(x,t)] \tag{6.11}$$

in which the diffusion operator on the left is balanced by a reaction-term $R[m(x,t)]$. Here, $R[m(x,t)]$ is written as the sum of two terms, one of which has a sigmoidal shape and corresponds to the interaction between z_2 and Mos (proportional to the concentration in z_2), the other of which is a linear degradation term. Thus we have

$$R[m(x,t)] = \Sigma[m(x,t)] - \gamma m(x,t), \tag{6.12}$$

with

$$\Sigma[m] = \frac{a\mathfrak{M}m^4}{\mathfrak{M}^4+\mathfrak{M}^2m^2+m^4}, \tag{6.13}$$

where a is the inverse of the characteristic time of creation of Mos, γ is the inverse of the characteristic time of its degradation and $\mathfrak{M}$ is the typical concentration. Recall that as it has been described previously, the expression of the reaction term $\Sigma[m]$ accounts for the feedback loop that drives the concentration of Mos depending on the MAPK concentration.

For both types of kinetics the Hill exponent is given by the product of the number of phosphorylation levels times cascade levels, hence four. For convenience in our mathematical treatment, we replaced both involved expressions by the simple expression (6.13), since we are interested chiefly in the role of ultrasensitivity, *i.e.* on the value of the Hill coefficient, on wave propagation. To this end, it suffices to sketch a qualitative shape, which can be inferred from Angeli's paper (see Figure 2). To be clear, from now on, we call the *Hill exponent* h the degree of the numerator and of the denominator (which are equal to obtain the sigmoidal shape) of the non–linearity (6.13). The reader has to keep in mind that is linked to the number of proteins involved in the cascade. The notion of *ultrasensitivity* refers to the steepness of the sigmoidal response curve which is characteristic of a signaling cascade: a very non linear response ensures an effective response to the signal received by the cell. The value $h = 4$ is typical for the MAPK cascade, but in principle of course other values $h > 1$ are equally possible. The limiting case $h = 1$ corresponds to a degenerate case, a non-cooperative cascade which is not of biological interest.

We are interested in the following questions:

i) Under what conditions does the system have three stationary states (two stable, one unstable)?

ii) Under what conditions does the system have propagating states, *i.e.*, the build-up of a profile connecting the two stable states and propagating from the cell surface to the nucleus?

iii) How does the propagation of the signal depend on the form of the non-linearity, and how does it depend on the chosen initial/boundary conditions?

iv) What is the selected propagation speed?

As we shall see in the next Section, by rendering the problem (7.16)–(6.13) dimensionless we extract a single significant parameter, denoted by α, which governs the dynamics. In Section 6.4.3, we will prove that there exists a constant α_0 such that for any $\alpha > \alpha_0$ there exist three stationary states : $0 < a_\alpha < b_\alpha$. The dependence of a_α and b_α on α is also explored. The stability of these stationary states is studied:

a) a_α is unstable (Lemma 6.4.11) and 0 is locally stable (Theorem 6.4.4);

b) for b_α no such precise result is obtained: we can exhibit sufficient condition for stability on the initial condition (Theorem 6.4.13) and on the boundary conditions (Theorem 6.4.14). In short, one has to inject sufficient mass into the system in order to reach the stationary state b_α. However, one has to point out that these conditions are not explicit enough to be numerically tractable.

Next, we investigate the existence of profiles connecting the stationary states: in Section 6.4.3, we prove the existence of a trajectory T_{c^*} connecting the two stationary states 0 and b_α, linked to an asymptotic speed of propagation c^* which depends only on the non–linearity. It is independent of either the initial condition or the boundary condition. Note that in our case, we do not have an explicit formula for the asymptotic velocity as a function of the parameters. Theorem 6.4.16 shows the existence of a wave front with velocity c^* connecting the two stationary states 0 and b_α provided the conditions evoked in b) are satisfied. Eventually, we end with numerical simulations in Section 6.5.1 exploring the behavior of the solution for several boundary and initial conditions. In Section 6.5.2 we compute numerically the asymptotic velocity c^* and describe its dependence on the parameter α.

6.3 Dimensionless equations

In order to complete the mathematical analysis of the problem, let us write the problem (7.16)–(6.13) in a dimensionless way. Note that the constants a and γ have the same dimension, that is, one over a time. In the same way, $\mathfrak{M}$ has the same dimension as m. As usual, the diffusion constant D is homogeneous to the square of a length divided by a time. We set now:

$$\bar{t} = \gamma t, \quad \bar{m} = \frac{m}{\mathfrak{M}}, \quad \bar{x} = \sqrt{\frac{\gamma}{D}}x, \quad \alpha = \frac{a}{\gamma}.$$

Omitting the overlines the system takes the following dimensionless form:

$$\partial_t m - \Delta m = z_\alpha(m) - m, \tag{6.14}$$

where

$$z_\alpha(m) = \frac{\alpha m^4}{m^4 + m^2 + 1}. \tag{6.15}$$

Note that in this dimensionless form all the information is contained in the parameter α. Thus from now on, the discussion will mainly concern this parameter α. The problem (6.14) is set on $t \geq 0$ and $x \in [0, L]$. Biologically, L represents the distance between the cell membrane and the nuclear membrane. Hence the problem has to be completed by an initial condition:

$$m(0, x) = m_0(x) \quad \text{on} \quad [0, L] \tag{6.16}$$

and boundary conditions:

- At $x = 0$, we have
$$m(t, 0) = \psi(t). \tag{6.17}$$
- At $x = L < \infty$ we choose a free outgoing condition so that we impose $\partial_x m(x = L) = 0$ (Neumann condition).

However for the convenience of mathematical analysis we simply set $L = +\infty$.

6.4 Stability and wave propagation

From now on, we denote

$$f(m) = z_\alpha(m) - m. \tag{6.18}$$

6.4.1 Preliminary remarks

A plot of the two functions $m \mapsto z_\alpha(m)$ and $m \mapsto m$ reveals, under conditions to be detailed below, three intersections. These correspond to stationary states of the kinetics for homogeneous profiles m. Two of the states, for small and large values of m, are stable, while the intermediate state is unstable. The profile of $R[m(x,t)]$ thus corresponds to the reaction term of a bistable system.

If we plot the curve corresponding to the function $z : m \mapsto z_\alpha(m)$ and the line $y = m$, see Figure 6.5 we note that these two curves can have one or three distinct intersection points depending on the value of α.

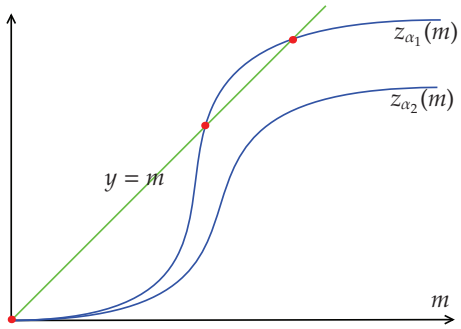


Figure 6.5: Functions z_{α_1} and z_{α_1} with $\alpha_1 > \alpha_2$ and the line $y = m$.

Let us now evaluate α_0 such that if $\alpha > \alpha_0$ then there exist three stationary points. Let us denote

$$f_\alpha(m) = \frac{m}{1+m^2+m^4} \underbrace{(\alpha m^3 - (1+m^2+m^4))}_{P_\alpha(m)},$$

thus

$$P'_\alpha(m) = 3\alpha m^2 - (2m + 4m^3) = m(-4m^2 + 3\alpha m - 2).$$

It is easy to see that if f has three non negative roots, then necessarily $\Delta = 9\alpha^2 - 32 > 0$. Thus, denoting $0 < m_1 < m_2$ the three non negative roots of $P'_\alpha(m)$, a sufficient condition to have three stationary points is $f_\alpha(m_2) > 0$. Numerically we find:

$$\alpha_0 = 2.463.$$

In case of $\alpha > \alpha_0$, f_α vanishes at 0, a_α and b_α. We plot the shape of the graph of f_α, in Figure 6.6.

We set

$$F_\alpha(x) := \int_0^x f_\alpha(u)\,\mathrm{d}u = \alpha x - \frac{x^2}{2} - \alpha\frac{1}{\sqrt{3}}\left(\arctan\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right) + \arctan\left(\frac{2}{\sqrt{3}}\left(x-\frac{1}{2}\right)\right)\right). \tag{6.19}$$

The quantity $-F_\alpha$ is naturally interpreted as a mechanical potential of the system. Figure 6.7 represents the shape of the potential in the three different cases $\alpha < \alpha_c$, $\alpha = \alpha_c$, $\alpha > \alpha_c$ and makes the stable states appear, where $\alpha_c > \alpha_0$ satisfies : $F_{\alpha_c} = \int_0^{b_{\alpha_c}} f_{\alpha_c}(u)\,\mathrm{d}u = 0$. Numerically we find :

$$\alpha_c = 2.878. \tag{6.20}$$

Furthermore, we impose

$$\int_0^{b_\alpha} f_\alpha(u)\,\mathrm{d}u > 0. \tag{6.21}$$

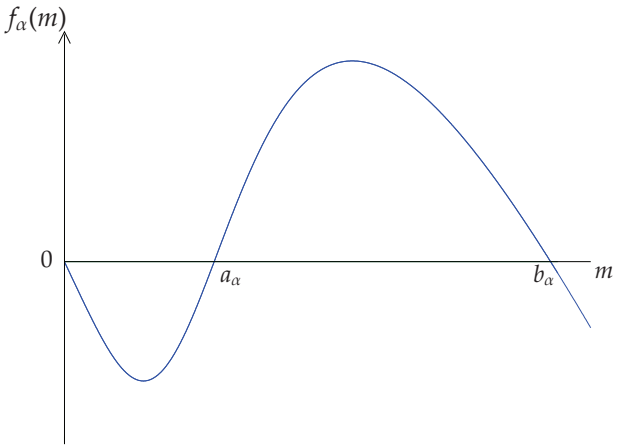
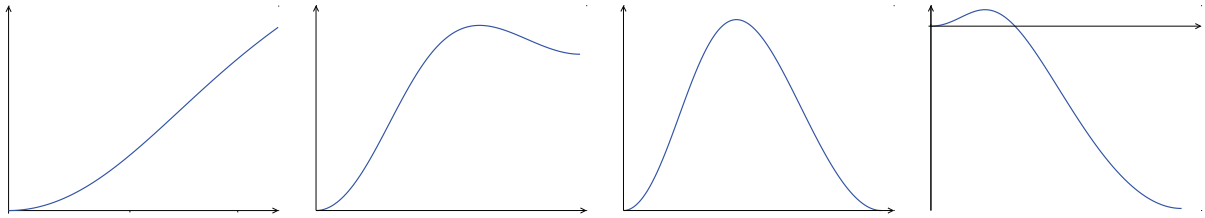


Figure 6.6: Function f_α.



(a) $\alpha = 2 < \alpha_0$ (b) $\alpha_0 < \alpha = 2.6 < \alpha_0$ (c) $\alpha = \alpha_c$ (d) $\alpha = 4 > \alpha_c$

Figure 6.7: Function $-F_\alpha$ for 4 different values of α

Referring to Figure 6.7, this condition means that the stable state b_α has negative and minimal energy ($-F_\alpha(b_\alpha) < 0$). Thanks to this condition we have:

$$\exists \kappa_\alpha \in (a_\alpha, b_\alpha) \text{ s.t. } F_\alpha(\kappa_\alpha) = \int_0^{\kappa_\alpha} f_\alpha(u)\,\mathrm{d}u = 0, \tag{6.22}$$

and then $\forall q \in (\kappa_\alpha, b_\alpha)$, $F_\alpha(q) > 0$ and $F'_\alpha(q) = f_\alpha(q) > 0$.

Thanks to Ferrari's formula we can give an explicit equivalent of the biggest root of f_α that we denote by b_α. After a simple computation we find out that

$$b_\alpha \sim \alpha, \tag{6.23}$$

and thus we get:

$$\kappa_\alpha \sim \frac{\alpha^2}{2}. \tag{6.24}$$

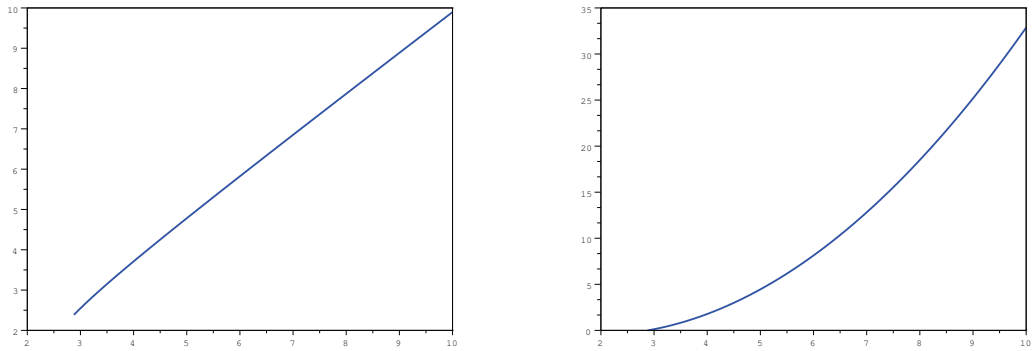
To illustrate we plot $\alpha \mapsto b_\alpha$ and $\alpha \mapsto F_\alpha(B_\alpha)$ in Figure 6.8. From this point on, we will omit the index α.

For understanding convenience, we prefer detailing first the mathematical analysis on the half–space $\mathbb{R} \times \mathbb{R}^+$. We adapt then the proofs to the case when we have a boundary.

6.4.2 Study of the pure initial value problem on the half–plane $\mathbb{R} \times \mathbb{R}^+$

The problem (6.14),(6.16) on the half–plane reads as follows:

$$\left\{ \begin{array}{rcll} m_t & = & \Delta m + f(m) & \text{in } \mathbb{R} \times \mathbb{R}^+ \\ m(x,0) & = & m_0(x) & \text{in } \mathbb{R}. \end{array} \right. \tag{6.25}$$



(a) $\alpha \mapsto b_\alpha$ (b) $\alpha \mapsto F_\alpha(b_\alpha)$

Figure 6.8: $\alpha \mapsto b_\alpha$ and $\alpha \mapsto F_\alpha(b_\alpha)$ for $\alpha_c < \alpha < 10$

A maximum principle.

Two lemmas are very useful to the study of our equation. They are given here without any proof, for more details see [AW75]. The first lemma is a version of the maximum principle.

Lemma 6.4.1 *Let $u : \mathbb{R} \times [0,T] \longrightarrow [0,b]$ and $v : \mathbb{R} \times [0,T] \longrightarrow [0,b]$ satisfy the inequalities*

$$\begin{array}{rll} u_t - \Delta u - f(u) \geq v_t - \Delta v - f(v) & \textit{in} & (x_1,x_2)\times(0,T] \\ 0 \leq v(x,0) \leq u(x,0) \leq b & \textit{in} & (x_1,x_2) \end{array}$$

where $-\infty \leq x_1 < x_2 \leq +\infty$ and $0 \leq T \leq \infty$. Moreover, if $x_1 > -\infty$, assume that

$$0 \leq v(x_1,t) \leq u(x_1,t) \leq b \quad \textit{on} \quad [0,T]$$

and if $x_2 < \infty$ assume that

$$0 \leq v(x_2,t) \leq u(x_2,t) \leq b \quad \textit{on} \quad [0,T] \ .$$

Then $u \geq v$, and if $u(x,0) > v(x,0)$ in an open sub–interval of (x_1,x_2) then $u > v$, in $(x_1,x_2)\times(0,T]$.

Lemma 6.4.2 *Let $q : \mathbb{R} \longrightarrow [0,b]$ be a solution of the ordinary differential equation*

$$q'' + f(q) = 0 \quad \textit{in} \quad (x_1,x_2)$$

where $-\infty \leq x_1 < x_2 \leq \infty$. If $x_1 > -\infty$ assume that $q(x_1) = 0$ and if $x_2 < \infty$ assume that $q(x_2) = 0$. Let $v(x,t)$ denote the solution of the initial value problem:

$$v_t = \Delta v + f(v) \quad \textit{in} \quad \mathbb{R}\times\mathbb{R}^+,$$

$$v(x,0) = \begin{cases} q(x) & \textit{in} \quad (x_1,x_2), \\ 0 & \textit{in} \quad \mathbb{R}\backslash(x_1,x_2). \end{cases}$$

Then $v(x,t)$ is a nondecreasing function of t for each x. Moreover,

$$\lim_{t\to\infty} v(x,t) = \tau(x)$$

uniformly in each bounded interval, where $\tau(x)$ is the smallest nonnegative solution of the ordinary differential equation

$$\tau'' + f(\tau) = 0$$

on the whole real line $\mathbb{R}$ that satisfies the inequality

$$\tau(x) \geq q(x) \quad \textit{in} \quad (x_1,x_2).$$

Stability and threshold results

In our case, because $f'(0) < 0$, we will see that we need a minimal amount of initial mass to reach the stable stationary point $x = b$. In this section, we give two results : the first one shows that if the initial condition is too small then the solution of (6.14) tends uniformly to zero as t tends to infinity. The second on the contrary shows that if initially we have enough mass then the solution tends to b as t tends to infinity.

The following lemma shows that the state $m \equiv a$ is unstable.

Lemma 6.4.3 *Let $m : \mathbb{R} \times \mathbb{R}^+ \longrightarrow [0, b]$ be a solution of (6.14) and $f(u) < 0$ in $(0, \gamma]$. If $m_0 : \mathbb{R} \longrightarrow [0, \gamma]$, then*

$$\lim_{t\to\infty} m(\cdot, t) = 0$$

uniformly on $\mathbb{R}$.

Proof. Let v be the solution of the initial value problem

$$\left\{ \begin{array}{rcl} v_t & = & \Delta v + f(v), \\ v(x,0) & = & \gamma. \end{array} \right.$$

Then, since the initial condition does not depend on x, v does not depend on x and satisfies the relation

$$t = \int_v^\gamma [-f(u)]^{-1} \, \mathrm{d}u.$$

Hence v goes to zero as $t \to \infty$ since $f(0) = 0$.

Since $v(x, 0) \geq m(x, 0)$, the Lemma follows from Lemma 6.4.1.

□

For any $\rho \in [0, a)$ we define

$$s(\rho) = \sup \left\{ \frac{f(u)}{u - \rho}, \; u \in (a, b) \right\}. \tag{6.26}$$

Notation 2 *In what follows we shall use the following usual notation:*

$$[u]^+ = \max(u, 0).$$

Theorem 6.4.4 *Let $m : \mathbb{R} \times [0, T] \longrightarrow [0, b]$ be the solution of equation (6.14). If for some $\rho \in [0, a)$*

$$\int_{-\infty}^{\infty} [m(x, 0) - \rho]^+ \, \mathrm{d}x < \sqrt{\frac{2\pi}{s(\rho)e}} (a - \rho)$$

then

$$\lim_{t\to\infty} m(x, t) = 0$$

uniformly on $\mathbb{R}$.

Proof. Fix $\rho \in [0, a)$.

Step 1: Construction of an upper–solution w. Let $w(x, t)$ the solution of

$$\begin{array}{rcl} w_t & = & \Delta w + s(\rho) w \\ w(x,0) & = & [m(x,0) - \rho]^+. \end{array}$$

Using lemma 6.4.1, we get $w \geq 0$ and then $w = [w]^+$. Since $s(\rho) > 0$ and $f(u) < 0$ in $[0, a)$ we have $f(m) \leq s(\rho)[m - \rho]^+$. Let

$$v(x, t) := m(x, t) - \rho,$$

then

$$\partial_t v - \Delta v - s[v]^+ \leq \partial_t m - \Delta m - f(m) = 0 = \partial_t w - \Delta w - s[w]^+.$$

Using again lemma 6.4.1 we get $v(x,t) \leq w(x,t)$.

Step 2: It exists η such that $w + \rho \leq \eta < a$. Moreover, $\partial_t(we^{-st}) - \Delta(we^{-st}) = 0$, thus, using the heat kernel:

$$\begin{aligned} w(x,t) &= \frac{1}{2\sqrt{\pi t}} e^{st} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{4t}} [m(\xi,0) - \rho]^+ \,\mathrm{d}\xi \\ &\leq \frac{1}{2\sqrt{\pi t}} e^{st} \int_{-\infty}^{\infty} [m(\xi,0) - \rho]^+ \,\mathrm{d}\xi \end{aligned}$$

Let us fix now $t = \dfrac{1}{2\,s(\rho)}$, we get:

$$w\left(x, \frac{1}{2\,s(\rho)}\right) \leq \sqrt{\frac{es(\rho)}{2\pi}} \int_{-\infty}^{\infty} [m(\xi,0) - \rho]^+ \,\mathrm{d}\xi =: \eta - \rho < a - \rho$$

and thanks to the hypothesis $\eta < a$.

We conclude using lemma 6.4.3.

□

Theorem 6.4.4 shows that the state $m \equiv 0$ is locally stable while Lemma 6.4.3 proves that the state $m \equiv a$ is unstable. In what follows, we show that $m \equiv 0$ is not globally stable.

Thanks to (6.21) we have:

$$\exists \kappa \in (a,b) \text{ s.t. } F(\kappa) = \int_0^{\kappa} f(u)\,\mathrm{d}u = 0, \tag{6.27}$$

and then $\forall q \in (\kappa, b)$, $F(q) > 0$ and $F'(q) = f(q) > 0$.

For all $\beta \in (\kappa, b)$ we define the length

$$\ell_\beta = 2 \int_0^{\beta} \frac{1}{\sqrt{2F(\beta) - F(u)}} \,\mathrm{d}u. \tag{6.28}$$

Finally, let q_β be the solution of

$$q_\beta'' + f(q_\beta) = 0, \tag{6.29}$$

with

$$q_\beta(0) = 0 \text{ and } q_\beta'(0) = \sqrt{2F(\beta)},$$

such that

$$\frac{1}{2} q_\beta'^2 + F(q_\beta) = F(\beta).$$

These conditions imply

$$q_\beta > 0,\ q_\beta(0) = q_\beta(\ell_\beta) = 0 \text{ and } q_\beta(x) \leq q_\beta\left(\frac{\ell_\beta}{2}\right) \text{ on } (0, \ell_\beta).$$

For more details see Appendix 6.8.

Theorem 6.4.5 *Let $m : \mathbb{R} \times [0,T] \longrightarrow [0,b]$ be the solution of the equation* (6.14)*. For some $\beta \in (\kappa, b)$ and some x_0 assume*

$$m(x,0) \geq q_\beta(x - x_0) \quad \textit{on} \quad (x_0, x_0 + \ell_\beta)$$

then

$$\lim_{t \to \infty} m(x,t) = b.$$

Proof.

Step 1: We apply Lemma 6.4.2 with $q(x) = q_\beta(x)$:
let define

$$v_t = \Delta v + f(v),$$

$$v(x,0) = \begin{cases} q_\beta(x - x_0) & \text{on} \quad (x_0, x_0 + \ell_\beta) \\ 0 & \text{on} \quad \mathbb{R}\backslash(x_0, x_0 + \ell_\beta). \end{cases}$$

$t \mapsto v(x,t)$ is a nondecreasing function and

$$\lim_{t\to\infty} v(x,t) = \tau(x),$$

uniformly on each bounded interval, where τ is the smallest nonnegative function of

$$\tau'' + f(\tau) = 0$$

on $\mathbb{R}$ which satisfies $\tau(x) \geq q_\beta(x - x_0)$ on $(x_0, +x_0 + \ell_\beta)$.

Step 2: We prove now that $\tau(x) \notin [\kappa, b)$, $\forall x \in \mathbb{R}$. Let us assume the contrary and get a contradiction. Suppose that

$$\exists x_0 \quad \text{s.t.} \quad \alpha = \tau(x_0) \in [\kappa, b)$$

then

$$\frac{1}{2}\tau'^2 + F(\tau) = k \geq F(\alpha),$$

since $\tau'^2 \geq 0$. Thus $\frac{1}{\sqrt{k-F(u)}}$ is integrable on $[0, \alpha]$ and $\tau(x)$ is implicitly defined by the equation

$$x = x_0 \pm \int_\tau^\alpha \frac{\mathrm{d}u}{\sqrt{2(k - F(u))}},$$

where the sign is determined by the sign of $\tau'(x_0)$. Hence $\tau(x)$ vanishes with $\tau' \neq 0$ at a finite value of x. Then τ can not be nonnegative and we get a contradiction.

Step 3: Note that

$$\tau\left(x_0 + \frac{\ell_\beta}{2}\right) \geq q_\beta\left(\frac{\ell_\beta}{2}\right) = \beta > \kappa,$$

then by continuity and because τ is the minimal solution, we get

$$\tau(x) \equiv b,$$

which proves the theorem.

□

Theorem 6.4.5 shows not only that the state $m \equiv 0$ is unstable under compactly perturbations, but that $m \equiv b$ is stable under such disturbances. Both Theorem 6.4.4 and 6.4.5 bring to light a threshold phenomenon: a disturbance of bounded support of the state $m \equiv 0$ which is sufficiently large on a sufficiently large interval grows to b, while a not sufficiently large perturbation tends to vanish. We will find again this behavior for the system written on the quarter–space $\mathbb{R}^+ \times \mathbb{R}^+$ with a boundary at $x = 0$.

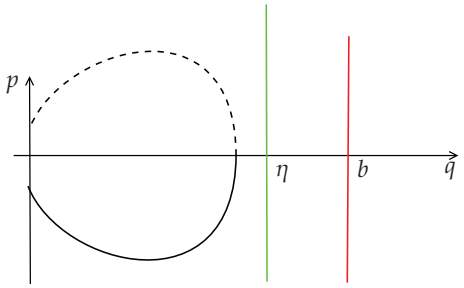


Figure 6.9: Trajectories for small value of c. The upper line is deduced by symetry.

Propagation

In this section we look at solutions of (6.25) as wave fronts. To this end we introduce the coordinate

$$\xi := x - ct, \quad \text{with} \quad c > 0, \tag{6.30}$$

and the new function

$$v(\xi, t) := m(\xi + ct, t). \tag{6.31}$$

Then, equation (6.25) becomes

$$v_t = \Delta_\xi v + cv_\xi + f(v). \tag{6.32}$$

Note that all the two Lemmas 6.4.1 and 6.4.2 can be extended to this equation.

We are interested in the steady–states equation corresponding to equation (6.47), which is

$$q'' + cq' + f(q) = 0, \tag{6.33}$$

that can be written as a first order system

$$\begin{array}{rcl} q' & = & p \\ p' & = & -cp - f(q). \end{array} \tag{6.34}$$

The functions $p(\xi)$ and $q(\xi)$ corresponding to a solution of (6.34) give a trajectory in the phase plane. Such a trajectory satisfies

$$\frac{\mathrm{d}p}{\mathrm{d}q} = -c - \frac{f(q)}{p}, \quad \text{for} \quad p \neq 0. \tag{6.35}$$

In the case $c = 0$. Each trajectory in the phase plane satisfies an equation of the form

$$\frac{1}{2}p^2 + F(q) = C.$$

Let fix η, $\kappa < \eta < b$. We have $F(\eta) > 0$. Then for any ε, such that $0 < \varepsilon < \sqrt{2F(\eta)}$, the trajectory through $(0, -\varepsilon)$ satisfies

$$F(\eta) > \frac{1}{2}\varepsilon^2 + F(0) = \frac{1}{2}p^2 + F(q),$$

which implies that the trajectory is confined in the strip $0 \leq q < \eta$. Hence by symmetry, the trajectory crosses the $p = 0$ axis at $(\alpha, 0)$ and the $q = 0$ axis at $(0, \varepsilon)$. By continuity this holds true for small velocities : there exists $\tilde{c}(\varepsilon)$ such that for all $c \in [0, \tilde{c}(\varepsilon))$ the trajectory crosses the $p = 0$ axis at $(\alpha, 0)$ and the $q = 0$ axis at $(0, \bar{\varepsilon})$ with $\bar{\varepsilon} > 0$. See Figure 6.9.

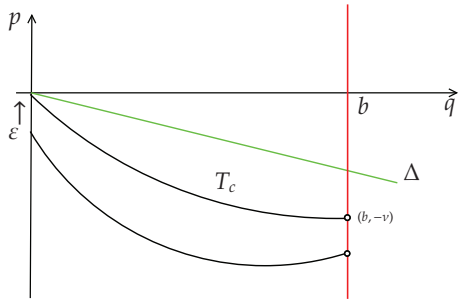


Figure 6.10: Trajectories and the line Δ defined at (6.36) for c large enough.

In the case $c > 0$**.** Since in our case we always have $c^2 > 4f'(0)$ the eigenvalues of the jacobian matrix of System (6.34) are real, there is no spiral point in $(0,0)$ and there exists one non trivial trajectory through $(0,0)$. In fact, if we seek a linear solution around the origin $p = \alpha q$ we find out that α should satisfy $\alpha^2 + c\alpha + f'(0) = 0$ which has a real solution if $c^2 > 4f'(0)$ which holds here since $f'(0) < 0$. The unique trajectory through $(0,-\varepsilon)$ with $\varepsilon > 0$, can not cross the trajectory that goes to the origin (Cauchy property). Hence if we take the limit trajectory as ν decreases to 0 we obtain a non trivial trajectory going to the origin. We denote by T_c this extremal trajectory.

We define

$$\sigma := \sup_{u\in[0,b]} \frac{f(u)}{u}.$$

Note that $\sigma \geq f'(0)$ and $f(u) \leq \sigma u$ for $u \in [0,b]$. It follows that for any trajectory T,

$$\frac{\mathrm{d}p}{\mathrm{d}q} \leq -c - \sigma\frac{q}{p}$$

at any point of T where $q \in [0,b]$ and $p < 0$. Furthermore

$$\Delta : \quad p = -\frac{1}{2}\left(c + \sqrt{c^2 - 4\sigma}\right)q \tag{6.36}$$

is a solution of

$$\frac{\mathrm{d}p}{\mathrm{d}q} = -c - \sigma\frac{q}{p},$$

and thus the trajectory through $(0,-\varepsilon)$, with $\varepsilon > 0$ lies under the line defined by (6.36). As ε tends to 0 we note that T_c is bounded above by the line defined by (6.36) and in particular connects the origin to a point of the shape $(b,-\nu)$ with $\nu > 0$. See Figure 6.10. Thus,

$$c^* := \inf\{c \geq 0 : \ \exists \nu > 0 \text{ s.t. } (b,-\nu) \in T_c\} \tag{6.37}$$

is well defined. In what follows, we show that c^* is the asymptotic speed of propagation associated to Equation (6.14) and that the limit trajectory T_{c^*} connects the origin to $(b,0)$.

Theorem 6.4.6 *Let $m(x,t) \in [0,b]$ be a solution of equation* (6.14) *in $\mathbb{R}\times\mathbb{R}^+$. If for some x_0*

$$m(x,0) = 0 \quad \textit{in} \quad (x_0,\infty),$$

then for each ξ and each $c > c^*$,

$$\lim_{t\to\infty} m(\xi + ct, t) = 0.$$

Proof. Step 1: Let $q_c(x)$ denote the solution of the steady state equation (6.33) in $\mathbb{R}^+$ which corresponds to the trajectory T_c. Since $q_c' = p$ (see (6.34)) and p is nonpositive along T_c, q_c is decreasing and has a limit as x tends to

zero which is necessarily 0. Let us define $w = b - m$. To prove the result we are going to show that the limit of w as t tends to infinity is larger than b.

The function w satisfies

$$w_t = \Delta w - f(b - w).$$

We apply the extension of Lemma 6.4.1 and 6.4.2 with $q = b - q_c(x - x_0)$ in (x_0, ∞). First Lemma 6.4.1 shows that

$$b - q_c(\xi - x_0) \leq w(\xi + ct, t) \quad \text{in} \quad (x_0, \infty) \times \mathbb{R}^+.$$

Let us define then the function v such that

$$v_t = \Delta_\xi v + c v_\xi + f(v)$$

and

$$b - v(\xi, 0) = \begin{cases} b - q_c(\xi - x_0) & \text{in} \quad (x_0, \infty) \\ 0 & \text{in} \quad \mathbb{R}\backslash(x_0, \infty) \end{cases}$$

Again using Lemma 6.4.1, we get

$$b - v(\xi, t) \leq w(\xi + ct, t) \quad \text{in} \quad (x_0, \infty) \times \mathbb{R}^+.$$

Futhermore, Lemma 6.4.2 shows that

$$\lim_{t \to \infty} v(\xi, t) = \tau(\xi),$$

where $b - \tau(\xi)$ is the smallest nonnegative function of

$$\tau'' + c\tau' + f(\tau) = 0 \quad \text{in} \quad \mathbb{R},$$

which satifies the inequality

$$b - \tau(\xi) \geq b - q_c(\xi - x_0) \quad \text{in} \quad (x_0, \infty). \tag{6.38}$$

Finally, it proves that

$$\liminf_{t \to \infty} (b - m(\xi + ct, t)) \geq b - \tau(\xi). \tag{6.39}$$

Step 2: We must show now that $\tau(\xi) \equiv 0$. For any $c > 0$ the trajectory T_c has slope s^- at the origin, where

$$s^- = \frac{1}{2}\left(-c - \sqrt{c^2 - 4f'(0)}\right).$$

Moreover, T_c is the unique trajectory with this slope at the origin and any other trajectory which approaches the origin with $q > 0$ must have the slope

$$s^+ = \frac{1}{2}\left(-c + \sqrt{c^2 - 4f'(0)}\right)$$

at the origin.

Since $c > c^*$, the trajectory T_c stays in the half–plane $p < 0$ for $q \in (0, b]$ and contains a point $(b, -\nu)$ with $\nu > 0$. Suppose that $\tau(\xi) \not\equiv 0$. Then the corresponding trajectory has either the slope s^+ or s^- at the origin. But it follows from (6.38) that the trajectory T should stay on the left of T_c in particular at the origin. Since T_c has the slope s^-, we deduce that necessarily T has also the slope s^- at $(0, 0)$ and by uniqueness $T \equiv T_c$. This leads us to a contradiction because there exists $\zeta \in \mathbb{R}$ such that $\tau(\zeta) = b$ and $\tau'(\zeta) = -\nu < 0$. Hence $\tau(\xi) > b$ for some $\xi < \zeta$ and this contradicts the nonnegativeness of the function $b - \tau$. We conclude that $\tau \equiv 0$.

Step 3: Equation (6.39) and the positiveness of m prove the theorem.

□

The following theorem gives the existence of a wave solution with velocity c^*.

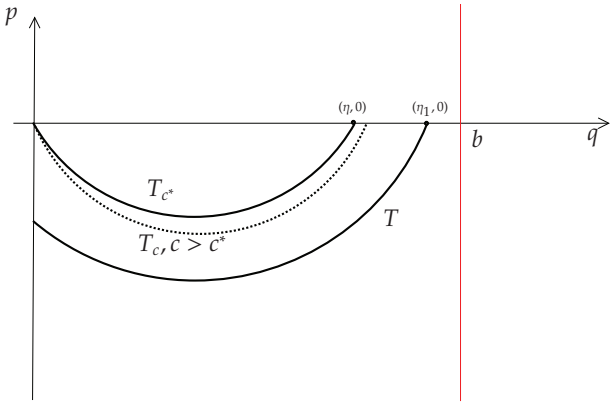


Figure 6.11: If $f(\eta) \neq 0$

Theorem 6.4.7 *The trajectory corresponding to the asymptotic speed of propagation, T_{c^*}, goes from $(b, 0)$ to $(0, 0)$ in the lower phase plane.*

Proof. We have already shown that the trajectory T_{c^*} exists and lies in the half–strip $q \in [0, b]$, $p \leq 0$ at least in a neighbourhood of the origin. T_{c^*} is minimal in the sense that there is no other trajectory which reaches the origin and lies below T_{c^*}.

Step 1: We first show that T_{c^*} must cross the $(p = 0)$ axis at some point η with $f(\eta) = 0$. Suppose that T_{c^*} does not cross the $(p = 0)$ axis, then it holds true for slightly smaller values of c which contradicts the definition of c^*. Hence T_{c^*} crosses the $(p = 0)$ axis at some point $(\eta, 0)$ with $\eta \in (0, b]$. Suppose $f(\eta) \neq 0$ then necessarily $f(\eta) > 0$: in fact T_{c^*} must go in the negative q–direction as $p < 0$ (see (6.34)) which implies that $-c - \dfrac{f(q)}{p}$ should be positive as (q, p) tends to $(\eta, 0)$ which is true only if $f(\eta) > 0$. Then there exists η_1 such that $f(\eta_1) > 0$ and $\eta_1 > \eta$. The part of the trajectory T, that goes through $(\eta_1, 0)$ lies below T_{c^*} and crosses the $(q = 0)$ axis at $(0, -v)$ with $v > 0$. By continuity, T_c for c slightly greater than c^*, will be bounded from below by T and can not cross the half line $q = b$, $p < 0$ which contradicts the definition of c^*. (See Figure 6.11.) We conclude that T_{c^*} hits the axis $(p = 0)$ at a point $(\eta, 0)$ where $f(\eta) = 0$.

Step 2: Let us show now that $\eta = b$. Since the trajectory nearby the origin lies in the half–plane $p < 0$ and, according to (6.35), if $f(q) < 0$ then the slope of $T_{c^*}$is negative and the trajectory can not cross the $(p = 0)$ axis at a zero of f which is the right end point of an interval where f is negative. Thus T_{c^*} must hit the $(p = 0)$ axis at $(b, 0)$.

□

Theorem 6.4.8 *Let $m : \mathbb{R} \times [0, T] \longrightarrow [0, b]$ be a solution of* (6.25). *Suppose that $\lim_{t\to\infty} m(x, t) = b$. Then for every c with $|c| \leq c^*$ and each ξ*

$$\lim_{t\to\infty} m(\xi + ct, t) = b.$$

Proof. Step 1: There exists a trajectory T that goes from a point on the positive p–axis to a point $(\beta, 0)$ and from there to a point on the negative p–axis.

Theorem 6.4.7 shows that T_{c^*} goes from $(b, 0)$ to $(0, 0)$ in the lower phase plane. Consider any $c < c^*$. Because of equation (6.35), the trajectory T_c lies above T_{c^*} and crosses the q–axis at a point $(\eta, 0)$, with $\eta \in (0, b)$. Then if $\beta \in (\kappa, b)$ (thus $\int_0^\beta f > 0$), with κ defined in (6.27), the lower part of the trajectory T through $(\beta, 0)$ stays below T_c. Therefore, T goes to the negative p–axis. Since $f(u) > 0$ nearby $u = \beta$, the slope of T is negative in the upper half–plane in a neighborhood of $(\beta, 0)$. Moreover, the slope is bounded below when p is bounded away from zero.

Therefore the trajectory does not go to infinity. From equation (6.35), we get:

$$\frac{\mathrm{d}p}{\mathrm{d}q} \leq -\frac{f(q)}{p}$$

which leads to the following inequality:

$$\int_{q_0}^{\beta} f(q)\,\mathrm{d}q \leq p^2(q_0), \quad \forall q_0 \in (0,\beta),\ p > 0.$$

As a consequence the trajectory T can not cross again the q–axis and necessarily it crosses the positive p–axis.

Step 2 : We have shown that for each $c \in (0, c^*)$ there is a trajectory T which connects the positive p–axis to the negative p–axis. T crosses the q–axis at a point $(\beta, 0)$ with $\beta \in (\kappa, b)$, and lies in the strip $q \in [0, \beta]$. Let q_β be the corresponding solution of $q'' + cq' + f(q) = 0$ for which $q_\beta(0) = 0$, $q'_\beta(0) > 0$. This solution is positive on a finite interval $(0, \delta)$ and $q_\beta(\delta) = 0$. Moreover, $q_\beta(x) \leq \beta < b$. Since we suppose that $\lim_{t\to\infty} m(x,t) = b$, uniformly on each interval, there is a time θ so that

$$m(x,\theta) \geq \beta \geq q_\beta(x) \quad \text{on} \quad [0,\delta].$$

Recalling that $v(\xi, t) = m(\xi + ct, t)$, the previous property can be written as follows:

$$v(\xi,\theta) \geq \beta \geq q_\beta(\xi) \quad \text{on} \quad [-c\theta, \delta - c\theta].$$

We apply the extention of Lemma 6.4.2 with $q(\xi) = q_\beta(\xi)$: let define

$$\bar{v}_t = \Delta\bar{v} + c\bar{v}_\xi + f(\bar{v}),$$

$$\bar{v}(\xi,\theta) = \begin{cases} q_\beta(\xi) & \text{on} \quad (-c\theta, \delta - c\theta) \\ 0 & \text{on} \quad \mathbb{R}\backslash(-c\theta, \delta - c\theta). \end{cases}$$

Then

$$\lim_{t\to\infty} \bar{v}(\xi, t) = \tau(\xi),$$

uniformly on each bounded interval, where τ is the smallest nonnegative function of

$$\tau'' + c\tau' + f(\tau) = 0$$

on $\mathbb{R}$ which satisfies $\tau(\xi) \geq q_\beta(\xi)$ on $(-c\theta, \delta - c\theta)$. Furthermore, note that according to the extension of Lemma 6.4.1 we have

$$v(\xi, t) \geq \bar{v}(\xi, t).$$

Step 3 : We show now that $\tau \equiv b$. Suppose that the trajectory corresponding to τ, designed by T_τ crosses the q–axis at one point $(\eta_\tau, 0)$, with $\beta \leq \eta_\tau \leq b$. Then similar arguments as in Step 1 show that T_τ connects the positive p–axis to the negative p–axis. Thus there exists ξ_0 such that $\tau(\xi_0) = 0$ and $\tau'(\xi_0) > 0$ which is in contradiction with τ nonnegative. Thus, $\tau \equiv b$, and since we suppose $m(x,t) \leq b$ we end up with:

$$\lim_{t\to\infty} m(\xi + ct, t) = \lim_{t\to\infty} v(\xi, t) = b.$$

□

6.4.3 Study of the initial-boundary value problem (6.14)–(6.16) in the quarter–space $\mathbb{R}^+ \times \mathbb{R}^+$

The problem (6.14)–(6.16) is rewritten just below:

$$\begin{cases} m_t = \Delta m + f(m) & \text{in} \quad \mathbb{R}^+ \times \mathbb{R}^+, \\ m(x,0) = m_0(x) & \text{in} \quad \mathbb{R}^+, \\ m(0,t) = \psi(t) & \text{in} \quad \mathbb{R}^+, \end{cases} \tag{6.40}$$

where ψ is a given function, $0 \leq \psi \leq b$. To be consistent we suppose

$$\psi(0) = m_0(0).$$

A maximum principle.

The proofs we made in the previous section were essentially based on the two Lemmas 6.4.1 and 6.4.2. To adapt the previous proofs to the problem (6.40) we give below an extension of Lemma 6.4.2 to the initial–boundary problem (6.40).

Lemma 6.4.9 *Let $q : \mathbb{R} \longrightarrow [0, b]$ be a solution of $q'' + f(q) = 0$ in (x_1, x_2) with $x_1 > 0$, and let $q(x_1) = q(x_2) = 0$.*
Let $v(x, t)$ denote the solution of the initial–boundary value problem

$$v_t = \Delta v + f(v) \quad in \quad \mathbb{R}^+ \times \mathbb{R}^+$$

$$v(x, 0) = \begin{cases} q(x) & in \quad (x_1, x_2) \\ 0 & in \quad \mathbb{R}^+ \backslash (x_1, x_2) \end{cases}$$

$$v(0, t) = \phi(t) \quad in \quad \mathbb{R}^+,$$

where ϕ is supposed nondecreasing and such that $\phi(0) = 0$ and $0 \leq \phi \leq b$.
Then v is nondecreasing with respect to t and

$$\lim_{t\to\infty} v(x, t) = \tau(x)$$

where τ is the smallest nonnegative solution of

$$\tau'' + f(\tau) = 0 \quad in \quad \mathbb{R}^+$$

which satisfies the inequalities

$$\tau(0) \geq \lim_{t\to\infty} \phi(t) \qquad and \qquad \tau(x) \geq q(x) \text{ in } (x_1, x_2).$$

Moreover, the convergence of v to τ is uniform on each closed bounded interval in the interior of $\mathbb{R}^+$.

Stability and threshold results

Again, since $f'(0) < 0$, we need a minimal amount of mass injected in the system either through the boundary to reach the stable stationary point $x = b$ or through the initial condition. In this section, we give two results : the first one shows that with no initial mass, if the boundary condition is too small in L^∞–norm then the solution of (6.40) tends uniformly to zero as t tends to infinity. The second shows that if the boundary condition combines a long enough support and a not too small value on this support then we observe a wave front. We have also as previously a condition on the initial condition that ensures the convergence to the steady state b.

For $\nu \in [0, \kappa)$ (where κ is defined at (6.22)), let us denote q_ν a solution of

$$q_\nu'' + f(q_\nu) = 0 \quad \text{in} \quad \mathbb{R}^+, \tag{6.41}$$

with $q_\nu(0) = \nu$.

In our case there is a solution of (6.41) such that

$$\lim_{x\to\infty} q_\nu(x) = 0.$$

In fact, choosing $q_\nu'(0) = -\sqrt{-2F(\nu)}$, where F is the primitive of f which vanishes at 0 (see (7.55)), one can prove using the first integral

$$\frac{1}{2} q_\nu'^2 + F(q_\nu) = 0$$

that the function converges to zero.

Theorem 6.4.10 *Let $m : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow [0, b]$ be the solution of the problem (6.40) with $m_0 \equiv 0$. If*

$$\nu = \sup_{t\in\mathbb{R}^+} \psi(t) < \kappa,$$

then $m(x, t) \leq q_\nu(x)$. In particular,

$$\lim_{x\to\infty} \limsup_{t\to\infty} m(x, t) = 0.$$

Proof. It is a consequence of Lemma 6.4.1.

□

The following Lemma shows that the stationary states a is unstable and 0 is locally stable. It is an extension of Lemma 6.4.3.

Lemma 6.4.11 *Let* $m : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow [0, b]$ *be a solution of* (6.40) *in* $\mathbb{R}^+ \times \mathbb{R}^+$ *and* $f(u) < 0$ *in* $(0, \gamma]$. *If* $m_0 : \mathbb{R}^+ \longrightarrow [0, \gamma]$ *and* $\sup_{\mathbb{R}^+} \psi(t) \leq \gamma$, *then*

$$\lim_{x\to\infty} \limsup_{t\to\infty} m(x,t) = 0.$$

Proof. Step1: Let w be solution of

$$\begin{array}{rcll} w_t & = & \Delta w + f(w) & \text{in} \quad \mathbb{R}^+ \times \mathbb{R}^+, \\ w(x,0) & = & \gamma & \text{in} \quad \mathbb{R}^+, \\ w(0,t) & = & \gamma & \text{in} \quad \mathbb{R}^+. \end{array}$$

Applying Lemma 6.4.1 we get $m(x,t) \leq w(x,t)$ in $\mathbb{R}^+ \times \mathbb{R}^+$.

Step 2: Since $f(0) = f(a) = 0$, Lemma 6.4.1 shows that $0 \leq w(x,t) \leq a$ in $\mathbb{R}^+ \times \mathbb{R}^+$. Furthermore note that $f \leq 0$ on $[0, a]$ thus

$$w_t - \Delta w \leq w_t - \Delta w - f(w) = 0 = u_t - \Delta u,$$

where $u \equiv \gamma$ in $\mathbb{R}^+ \times \mathbb{R}^+$. Applying again Lemma 6.4.1, we get $w(x,t) \leq \gamma$ in $\mathbb{R}^+ \times \mathbb{R}^+$.

Step 3: Note that on $[0, \gamma]$, $f(x) \leq -\left|\dfrac{f(\gamma)}{\gamma}\right| x =: -C_\gamma x$ with $C_\gamma > 0$. Let v be the solution of

$$\begin{array}{rcll} v_t & = & \Delta v - C_\gamma v & \text{in} \quad \mathbb{R}^+ \times \mathbb{R}^+, \\ v(x,0) & = & \gamma & \text{in} \quad \mathbb{R}^+, \\ v(0,t) & = & \gamma & \text{in} \quad \mathbb{R}^+. \end{array}$$

Using again Lemma 6.4.1, we show that $w(x,t) \leq v(x,t)$ thus $m(x,t) \leq v(x,t)$ in $\mathbb{R}^+ \times \mathbb{R}^+$.

Step 4: Finally, $\bar{v}(x,t) = e^{C_\gamma t}\left(v(x,t) - \gamma e^{-\sqrt{C_\gamma} x}\right)$ is the solution of

$$\begin{array}{rcll} \bar{v}_t & = & \Delta \bar{v} & \text{in} \quad \mathbb{R}^+ \times \mathbb{R}^+, \\ \bar{v}(x,0) & = & \gamma\left(1 - e^{-\sqrt{C_\gamma} x}\right) & \text{in} \quad \mathbb{R}^+, \\ \bar{v}(0,t) & = & 0 & \text{in} \quad \mathbb{R}^+. \end{array}$$

It is easy to see that $\bar{v}(x,t) \leq \gamma$ in $\mathbb{R}^+ \times \mathbb{R}^+$, thus $\lim_{t\to\infty} v(x,t) - \gamma e^{-\sqrt{C_\gamma} x} = 0$ and we conclude that

$$\lim_{x\to\infty} \limsup_{t\to\infty} m(x,t) = 0.$$

□

The following Theorem 6.4.12 is the extention of Theorem 6.4.4 when we add a boundary condition at $x = 0$. The proof is the same than for Theorem 6.4.4 : the treatment of the boundary condition with the heat kernel requires a bit more computations but the idea is exactly the same.

Theorem 6.4.12 *Let $m(x,t) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto [0,b]$ be the solution of* (6.40)*. Suppose that for some $T \in \mathbb{R}^+$ and some $\rho \in (0,a)$,*

$$\psi(t) \le \rho \quad in \quad (T,\infty) \tag{6.42}$$

and

$$e^{s(\rho)T}\left(\sqrt{\frac{e\,s(\rho)}{4\pi(s(\rho)T+1)}}\int_0^\infty [m_0(y)-\rho]^+\,\mathrm{d}y + s(\rho)\sqrt{\frac{e}{2\pi}}\int_0^T e^{-s(\rho)t}\,[\psi(t)-\rho]^+\,\mathrm{d}t\right) < a-\rho. \tag{6.43}$$

Then

$$\lim_{x\to\infty}\limsup_{t\to\infty} m(x,t) = 0.$$

Proof. Fix $\rho \in [0,a)$.

Step 1 : Construction of an upper–solution w. Let $w(x,t)$ be the solution of

$$\left\{\begin{array}{rcll} w_t &=& \Delta w + s(\rho)w & \text{in } \mathbb{R}^+\times\mathbb{R}^+,\\ w(x,0) &=& [m_0(x)-\rho]^+ & \text{in } \mathbb{R}^+,\\ w(0,t) &=& [\psi(t)-\rho]^+ & \text{in } \mathbb{R}^+. \end{array}\right.$$

Using Lemma 6.4.1 we get $w \ge 0$ and then $w = [w]^+$. Since $s(\rho) > 0$ and $f(u) < 0$ in $(0,a)$ we have $f(m) \le s(\rho)[m-\rho]^+$. Let $v(x,t) := m(x,t) - \rho$, then

$$v_t - \Delta v - s(\rho)[v]^+ \le m_t - \Delta m - f(m) = 0 = w_t - \Delta w - s(\rho)[w]^+.$$

Using Lemma 6.4.1 again we get $v(x,t) \le w(x,t)$.

Step 2 : There exists η such that $w + \rho \le \eta < a$. Let us define for $(x,t) \in \mathbb{R}^+\times\mathbb{R}^+$,

$$\bar{w}(x,t) := w(x,t)e^{-s(\rho)t}$$

and

$$W(x,t) = \bar{w}(x,t) - [\psi(t)-\rho]^+\,e^{-s(\rho)t}.$$

Then W satisfies

$$\left\{\begin{array}{rcll} W_t - \Delta W &=& -\partial_t\left([\psi(t)-\rho]^+\,e^{-s(\rho)t}\right) & \text{in } \mathbb{R}^+\times\mathbb{R}^+,\\ W(x,0) &=& [m_0(x)-\rho]^+ - [\psi(0)-\rho]^+ & \text{in } \mathbb{R}^+,\\ W(0,t) &=& 0 & \text{in } \mathbb{R}^+. \end{array}\right.$$

Let us now extend function W to $\mathbb{R}$ as an odd function:

$$\left\{\begin{array}{rcll} \bar{W}_t - \Delta\bar{W} &=& -\partial_t\left([\psi(t)-\rho]^+\,e^{-s(\rho)t}\right)sg(x) & \text{in } \mathbb{R}^+\times\mathbb{R},\\ \bar{W}(x,0) &=& \left([m_0(sg(x)x)-\rho]^+ - [\psi(0)-\rho]^+\right)sg(x) & \text{in } \mathbb{R},\\ \bar{W}(0,t) &=& 0 & \text{in } \mathbb{R}^+. \end{array}\right.$$

where

$$sg(x) = \left\{\begin{array}{rl} 1 & \text{if } x \ge 0\\ -1 & \text{if } x < 0 \end{array}\right.$$

Note that

$$\forall x \in \mathbb{R}^+, \quad \bar{W}(x) = W(x).$$

Omitting the overlines and using the heat kernel, the solution reads as follows:

$$W(x,t) = \int_{\mathbb{R}} \frac{e^{-\frac{|x-y|^2}{4t}}}{\sqrt{4\pi t}}\left([m_0(sg(y)y)-\rho]^+ - [\psi(0)-\rho]^+\right)sg(y)\,\mathrm{d}y$$
$$- \int_0^t\int_{\mathbb{R}} \frac{e^{-\frac{|x-y|^2}{4(t-\sigma)}}}{\sqrt{4\pi(t-\sigma)}}\partial_t\left([\psi(\sigma)-\rho]^+\,e^{-s(\rho)\sigma}\right)sg(y)\,\mathrm{d}y\,\mathrm{d}\sigma.$$

Let us denote $g(\sigma) = [\psi(\sigma) - \rho]^+ e^{-s(\rho)\sigma}$, and after an integration by parts, we get:

$$\begin{aligned} W(x,t) &= \int_{\mathbb{R}} \frac{e^{-\frac{|x-y|^2}{4t}}}{\sqrt{4\pi t}} [m_0(sg(y)y) - \rho]^+ sg(y)\,\mathrm{d}y + \int_0^t \int_{\mathbb{R}} \frac{e^{-\frac{|x-y|^2}{4(t-\sigma)}}}{8\sqrt{\pi}(t-\sigma)^{3/2}} \left(-\frac{|x-y|^2}{t-\sigma} + 2\right) sg(y)\,\mathrm{d}y\, g(\sigma)\,\mathrm{d}\sigma \\ &= I_1(t) + \int_0^t I_2(\sigma) g(\sigma)\,\mathrm{d}\sigma. \end{aligned}$$

Let us find an upperbound for I_1:

$$\begin{aligned} I_1(t) &= \int_{\mathbb{R}} \frac{e^{-\frac{|x-y|^2}{4t}}}{\sqrt{4\pi t}} [m_0(sg(y)y) - \rho]^+ sg(y)\,\mathrm{d}y \\ &= \int_0^\infty \frac{e^{-\frac{|x-y|^2}{4t}}}{\sqrt{4\pi t}} [m_0(y) - \rho]^+\,\mathrm{d}y - \int_{-\infty}^0 \frac{e^{-\frac{|x-y|^2}{4t}}}{\sqrt{4\pi t}} [m_0(-y) - \rho]^+\,\mathrm{d}y \\ &\leq \int_0^\infty \frac{e^{-\frac{|x-y|^2}{4t}}}{\sqrt{4\pi t}} [m_0(y) - \rho]^+\,\mathrm{d}y \\ &\leq \frac{1}{\sqrt{4\pi t}} \int_0^\infty [m_0(y) - \rho]^+\,\mathrm{d}y. \end{aligned}$$

In particular,

$$I_1\left(T + \frac{1}{s(\rho)}\right) \leq \sqrt{\frac{s(\rho)}{4\pi(s(\rho)T+1)}} \int_0^\infty [m_0(y) - \rho]^+\,\mathrm{d}y.$$

Let us compute the spatial integral I_2:

$$\begin{aligned} I_2 &= \int_{\mathbb{R}} \frac{e^{-\frac{|x-y|^2}{4(t-\sigma)}}}{8\sqrt{\pi}(t-\sigma)^{3/2}} \left(-\frac{|x-y|^2}{t-\sigma} + 2\right) sg(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^+} \frac{e^{-\frac{|x-y|^2}{4(t-\sigma)}}}{8\sqrt{\pi}(t-\sigma)^{3/2}} \left(-\frac{|x-y|^2}{t-\sigma} + 2\right) sg(y)\,\mathrm{d}y - \int_{\mathbb{R}^-} \frac{e^{-\frac{|x-y|^2}{4(t-\sigma)}}}{8\sqrt{\pi}(t-\sigma)^{3/2}} \left(-\frac{|x-y|^2}{t-\sigma} + 2\right) sg(y)\,\mathrm{d}y \\ &= \int_{\frac{-x}{2\sqrt{t-\sigma}}}^{\frac{x}{2\sqrt{t-\sigma}}} \frac{1}{2\sqrt{\pi}\sqrt{t-\sigma}} e^{-z^2}(4z^2+2)\,\mathrm{d}z \\ &= \left[\frac{ze^{-z^2}}{\sqrt{\pi}(t-\sigma)}\right]_{\frac{-x}{2\sqrt{t-\sigma}}}^{\frac{x}{2\sqrt{t-\sigma}}} \\ &\leq \frac{e^{-1/2}}{\sqrt{2\pi}(t-\sigma)}. \end{aligned}$$

Finally, thanks to (6.42) we get

$$\int_0^{T+\frac{1}{s(\rho)}} I_2(\sigma) g(\sigma)\,\mathrm{d}\sigma \leq \frac{s(\rho)}{\sqrt{2e\pi}} \int_0^T e^{-s(\rho)t} [\psi(t) - \rho]^+\,\mathrm{d}t.$$

To conclude,

$$W\left(x, T + \frac{1}{s(\rho)}\right) \leq \sqrt{\frac{s(\rho)}{4\pi(s(\rho)T+1)}} \int_0^\infty [m_0(y) - \rho]^+\,\mathrm{d}y + \frac{s(\rho)}{\sqrt{2e\pi}}) \int_0^T e^{-s(\rho)t} [\psi(t) - \rho]^+\,\mathrm{d}t,$$

which thanks to (6.43), leads to

$$\exists \eta < a, \quad s.t. \quad w(x,t) < \eta - \rho \quad (x,t) \in \mathbb{R}^+ \times \mathbb{R}^+.$$

We conclude using Lemma 6.4.11.

□

Theorem 6.4.13 *Let $m : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow [0,b]$ be the solution of* (6.40)*. For some $\beta \in (\kappa, b)$ and some $x_0 > 0$ assume*

$$m_0(x) \geq q_\beta(x - x_0) \quad on \quad (x_0, x_0 + \ell_\beta) \tag{6.44}$$

then

$$\lim_{x\to\infty} \lim_{t\to\infty} m(x,t) = b,$$

where ℓ_β and q_β are defined at (6.28) *and* (6.29)

Proof. The proof is exactly the same as the one of Theorem 6.4.5, using Lemma 6.4.9 instead of Lemma 6.4.2.

□

Theorem 6.4.14 *Let $m : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow [0,b]$ be the solution of* (6.40)*. Let κ defined at* (6.22)*. If for any $\beta \in (\kappa, b)$ there is a positive time T_β such that*

$$\psi(t) \geq \beta \quad on \quad (t_0, t_0 + T_\beta) \tag{6.45}$$

for some nonnegative t_0, then

$$\lim_{x\to\infty} \liminf_{t\to\infty} m(x,t) = b.$$

Proof. Step 1: Let χ be a smooth function such that

$$\chi(t) = \begin{cases} 0 & \text{on} \quad (-\infty, 0), \\ \beta & \text{on} \quad (1, \infty). \end{cases}$$

Let w denote the solution of

$$\begin{array}{lll} w_t = \Delta w + f(w) & \text{in} & \mathbb{R} \times \mathbb{R}^+, \\ w(x,0) = m_0(x) & \text{in} & \mathbb{R}^+, \\ w(0,t) = \chi(t) & \text{in} & \mathbb{R}^+. \end{array}$$

Using Lemma 6.4.9, we get

$$\lim_{t\to\infty} w(x,t) = \tau(x),$$

where τ is the smallest nonnegative solution of

$$\begin{cases} \tau'' + f(\tau) = 0, \\ \tau(0) = \beta \text{ and } \tau(x) \geq m_0(x) \text{ on } \mathbb{R}. \end{cases} \tag{6.46}$$

Step 2: Let us show that the problem (6.46) has only one solution since $\beta > \kappa$. The first integral gives us

$$\frac{1}{2}(\tau')^2 + F(\tau) = \frac{1}{2}(\tau'(0))^2 + F(\beta) = k \geq F(\beta) > 0.$$

First of all, τ can not vanish at some point : if $\tau(x_0) = 0$ then $(\tau'(x_0))^2 > 0$ which is in contradiction with τ being nonnegative. Assume next that there exists x_0 such that $\tau'(x_0) = 0$, then

$$\forall x \in \mathbb{R}, \quad \frac{1}{2}(\tau'(x))^2 = F(\tau(x_0)) - F(\tau(x)).$$

Furthermore $F(\tau(x_0)) > F(\beta)$ and F is increasing for $x > \beta$. This shows that $\tau' \leq 0$ on $[x_0 - h, x_0 + h]$. Finally, $\forall x \in \mathbb{R}$, $\tau' \leq 0$. This implies that $\tau'(x) = -\sqrt{k - F(\tau)} \leq -\sqrt{k - F(\beta)}$ and necessarily, τ vanishes at some point which is a contradiction. Finally $\tau' > 0$ and then τ is increasing.
τ is increasing and bounded by b, it has a limit which is necessarily b and according to the first integral we get

$$\tau'(0) = \sqrt{2(F(b) - F(\beta))}.$$

Then τ is uniquely determined and since $\tau' > 0$ we have $\tau(x) > \beta$ on $\mathbb{R}_*^+$.
Step 3: Recall the function q_β is defined at (6.29). Since the convergence of $w(x,t)$ towards $\tau(x)$ as t tends to infinity is uniform on each bounded interval, in particular, $w(x,t)$ converges to $\tau(x)$ uniformly on $[1, 1 + \ell_\beta]$, where ℓ_β is defined at (6.28).

Thus there exists a time T_β such that

$$w(x, T_\beta) \geq q_\beta(x - 1) \quad \text{on} \quad [1, \ell_\beta + 1].$$

Futhermore, thanks to (6.45) $\psi(t + t_0) \geq \chi(t)$ on $[0, T_\beta]$. Lemma 6.4.1 shows that

$$m(x, t + t_0) \geq w(x, t) \quad \text{on} \quad \mathbb{R}^+ \times [0, T_\beta],$$

and since

$$w(x, T_\beta) \geq q_\beta(x - 1) \quad \text{on} \quad [1, 1 + \ell_\beta],$$

we get

$$m(x, T_\beta + t_0) \geq q_\beta(x - 1) \quad \text{on} \quad (1, \ell_\beta + 1).$$

Step 4: Let u be the solution of

$$u_t = \Delta u + f(u) \quad \text{in} \quad \mathbb{R}^+ \times \mathbb{R}^+$$

$$u(x, 0) = \begin{cases} q_\beta(x - 1) & \text{in} \quad (1, 1 + \ell_\beta) \\ 0 & \text{elsewhere} \end{cases}$$

$$u(0, t) = \chi(t)$$

Lemma 6.4.9 shows that $\lim\limits_{t \to \infty} u(x,t) = \bar{\tau}(x)$ where $\bar{\tau}$ is the smallest nonnegative solution of $\tau'' + f(\tau) = 0$ which satisfies $\bar{\tau}(0) \geq lim_{t \to \infty} \chi(t) = \beta$ and $\bar{\tau}(x) > q_\beta(x-1)$ on $(1, 1 + \ell_\beta)$ and since $m(x, T_\beta + t_0) \geq q_\beta(x-1)$ in $(1, 1 + \ell_\beta)$, Lemma 6.4.1 shows that $m(x,t) \geq u(x,t)$.
Step 5: Finally,

$$\liminf_{t \to \infty} m(x,t) \geq \bar{\tau}(x)$$

which implies

$$b \geq \lim_{x \to \infty} \liminf_{t \to \infty} m(x,t) \geq \lim_{x \to \infty} \bar{\tau}(x)$$

and since

$$\bar{\tau}\left(\frac{1}{2}\ell_\beta + 1\right) \geq q_\beta\left(\frac{\ell_\beta}{2}\right) = \beta > \kappa$$

the same argument as at step 2 shows that $\bar{\tau}$ is increasing and tends to b as x tends to infinity. This allows to conclude that

$$\lim_{x \to \infty} \liminf_{t \to \infty} m(x,t) = b.$$

□

Theorems 6.4.13 and 6.4.14 show not only that the state $m \equiv 0$ is unstable under compactly supported perturbations, but that $m \equiv b$ is stable under such perturbations. Theorems 6.4.4, 6.4.13 and 6.4.14 bring to light a threshold phenomenon: a disturbance of bounded support of the state $m \equiv 0$ which is sufficiently large on a sufficiently large interval grows to b, while a not sufficiently large perturbation tends to vanish. The reader has to note that sufficient conditions to have the convergence to the steady state b are given on the initial condition or on the boundary condition. One can imagine that if we have a boundary condition and an initial condition which do not fulfill the conditions (6.43), (6.45) and (6.44), with the initial condition 'close enough' to the positive boundary we can observe a convergence to the steady state b, where 'close enough' has to be understood in the sense that we need enough mass in a small area at the same time. This will be investigated numerically at Section 6.5.1.

Propagation

In this section we look at solutions of (6.40) as wave fronts. We use again the moving coordinates introduced at (6.30)–(6.31). Then the set $\mathbb{R}^+ \times \mathbb{R}^+$ is mapped onto the set $\{(\xi, t) : \xi > -ct,\ t > 0\}$. Equation (6.40) becomes

$$v_t = \Delta_\xi v + c v_\xi + f(v). \tag{6.47}$$

Note that both Lemmas 6.4.1 and 6.4.9 can be extended to this equation. In particular Lemma 6.4.9 is extended as follows.

Lemma 6.4.15 *Let $q : \mathbb{R}^+ \longrightarrow [0, b]$ be a solution of $q'' + cq + f(q) = 0$ in (ξ_1, ξ_2) with $\xi_1 > 0$, and let $q(\xi_1) = q(\xi_2) = 0$.*

Let $v(\xi, t)$ denote the solution of the initial–boundary value problem

$$v_t = \Delta v + c v_\xi + f(v) \quad in \quad \{(\xi, t) : \xi > -ct,\ t > 0\},$$

$$v(\xi, 0) = \begin{cases} q(\xi) & in \quad (\xi_1, \xi_2), \\ 0 & in \quad \{(\xi, t) : \xi > -ct,\ t > 0\} \setminus (\xi_1, \xi_2), \end{cases}$$

$$v(-ct, t) = 0 \quad in \quad \mathbb{R}^+.$$

Then v is nondecreasing in t and

$$\lim_{t \to \infty} v(\xi, t) = \tau(\xi)$$

where τ is the smallest nonnegative solution of

$$\tau'' + c\tau' + f(\tau) = 0 \quad in \quad \mathbb{R}^+$$

which satisfies the inequality

$$\tau(\xi) \geq q(\xi) \quad in \quad (\xi_1, \xi_2)$$

Theorem 6.4.16 *Let $m(x, t) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto [0, b]$ be the solution of* (6.40)*. The asymptotic velocity c^* is defined as in* (6.37)*.*

a) *If for some x_0, $m(x, 0) = 0$ in (x_0, ∞). Then for any $c > c^*$ and any real ξ*

$$\lim_{t \to \infty} m(\xi + ct, t) = 0.$$

b) *If $\lim_{x \to \infty} \liminf_{t \to \infty} m(x, t) = b$, then for any $c \in (0, c^*)$ and any real ξ*

$$\lim_{t \to \infty} m(\xi + ct, t) = b.$$

Proof. Point a) is proved in exactly the same manner as Theorem 6.4.6 using Lemma 6.4.15 instead of Lemma 6.4.2. The proof of point b) is the same as the proof of Theorem 6.4.8.

□

6.5 Numerical Study of the Wave Propagation

In this last part of the paper we describe numerical simulations of the system at Section 6.5.1 exploring the behavior of the solution for several boundary and initial conditions. In Section 6.5.2 we compute numerically the asymptotic velocity c^*, and describe its dependence on the parameter α.

6.5.1 Approximation of the solution

We compute the solution using a Crank–Nicolson scheme. It is a scheme of order 2 in space and time with no stability condition. Let us fix in this section $\alpha = 4 > \alpha_c$ defined at (6.20).

Following Theorem 6.4.10 the appearance of a front should depend on the value of $\sup_t \psi(t)$. The value of κ, which itself depends on α as defined in (6.22), has a significant influence on the behavior of the system. In Figure 6.12, we plot the graph of the function $\kappa : \alpha \mapsto \kappa(\alpha)$. We observe that this function is decreasing: κ increases with the degradation term (which is proportional to $1/\alpha$). This is relevant biologically : the amount of mass needed to observe a wave front is reduced when the degradation rate goes down. For the simulation of the wave front formation, we choose $\alpha = 4$, which leads to $\kappa = 1.24$. In Figure 6.13, we illustrate the threshold phenomena stated in Theorem 6.4.10.

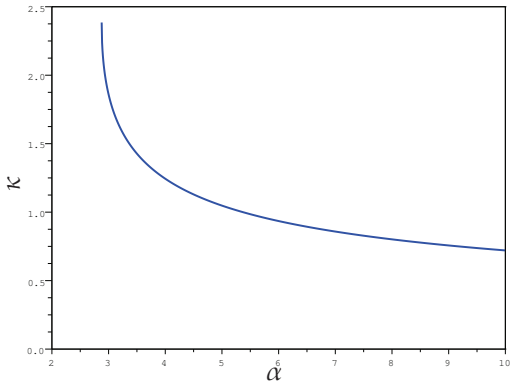
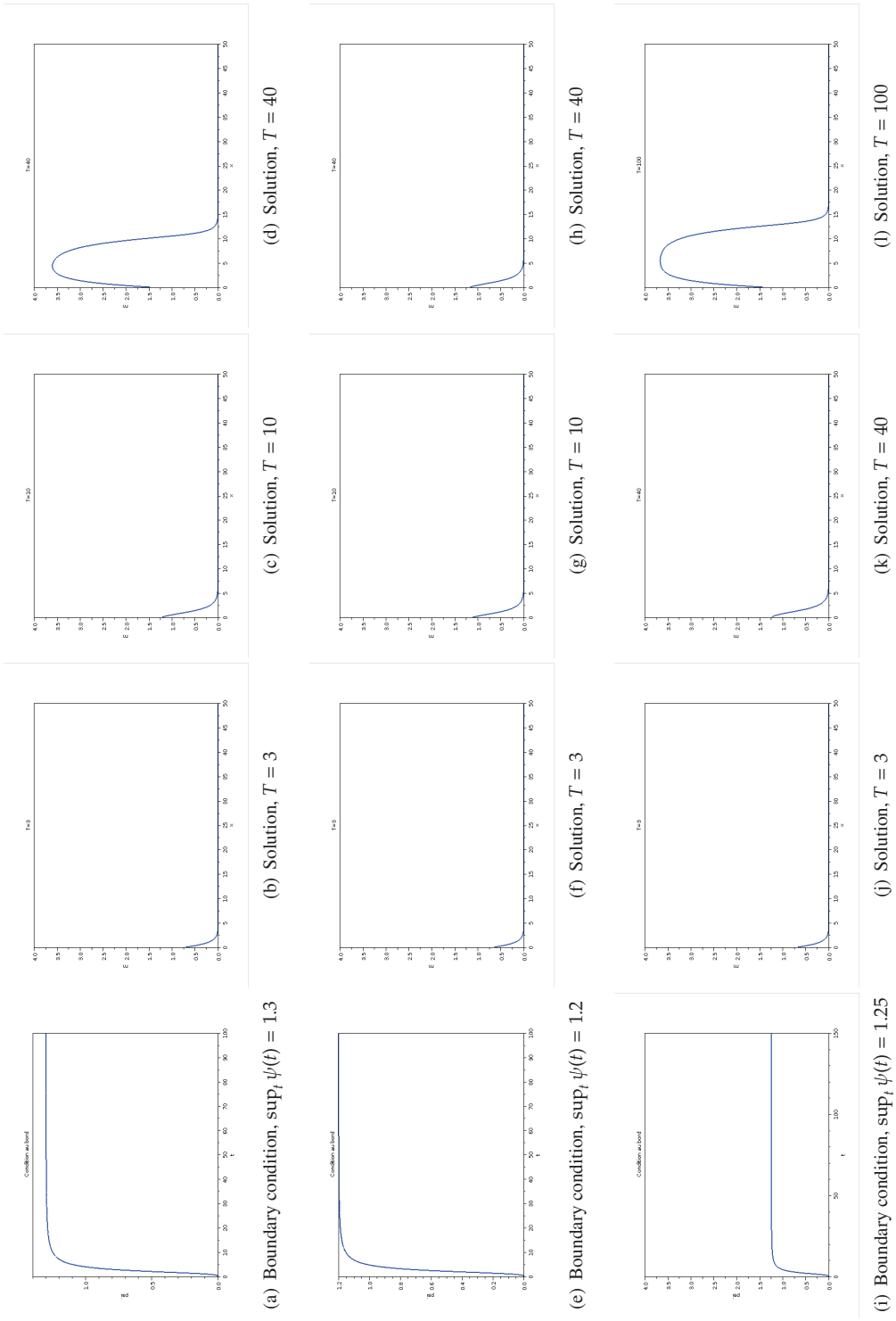


Figure 6.12: $\kappa : \alpha \mapsto \kappa(\alpha)$

Now Theorem 6.4.14 suggests that even when the hypothesis $\sup_t \psi(t) > \kappa$ is fulfilled, the support of ψ must be long enough. This is confirmed numerically (see Figure 6.14).

Theorem 6.4.13 states that in case of an homogeneous Dirichlet boundary condition, the initial condition has to be large enough to ensure the appearance of a wave front. In Figure 6.15, we study the influence of the initial condition. Note that we choose a gaussian initial profile, because from our point of view it is biologically relevant. We observe that the output, *i.e.* the existence of a wave front or not, should depend on the initial mass. This is not totally true, applying some different shape for a constant initial mass we observe different behaviors : the shape of the initial conditions has an influence on the output of the cascade. (See Figures 6.16 and 6.17.)

Now, we combine both positive boundary and initial conditions. (See Figure 6.18.) The reader has to note that there is a delay due to the time needed by the mass located in the boundary to reach the initial condition (compare for instance the two first line of Figure 6.18). This illustrates in particular our comment on the idea to be 'close enough' at the end of Section 6.4.3, page 169.

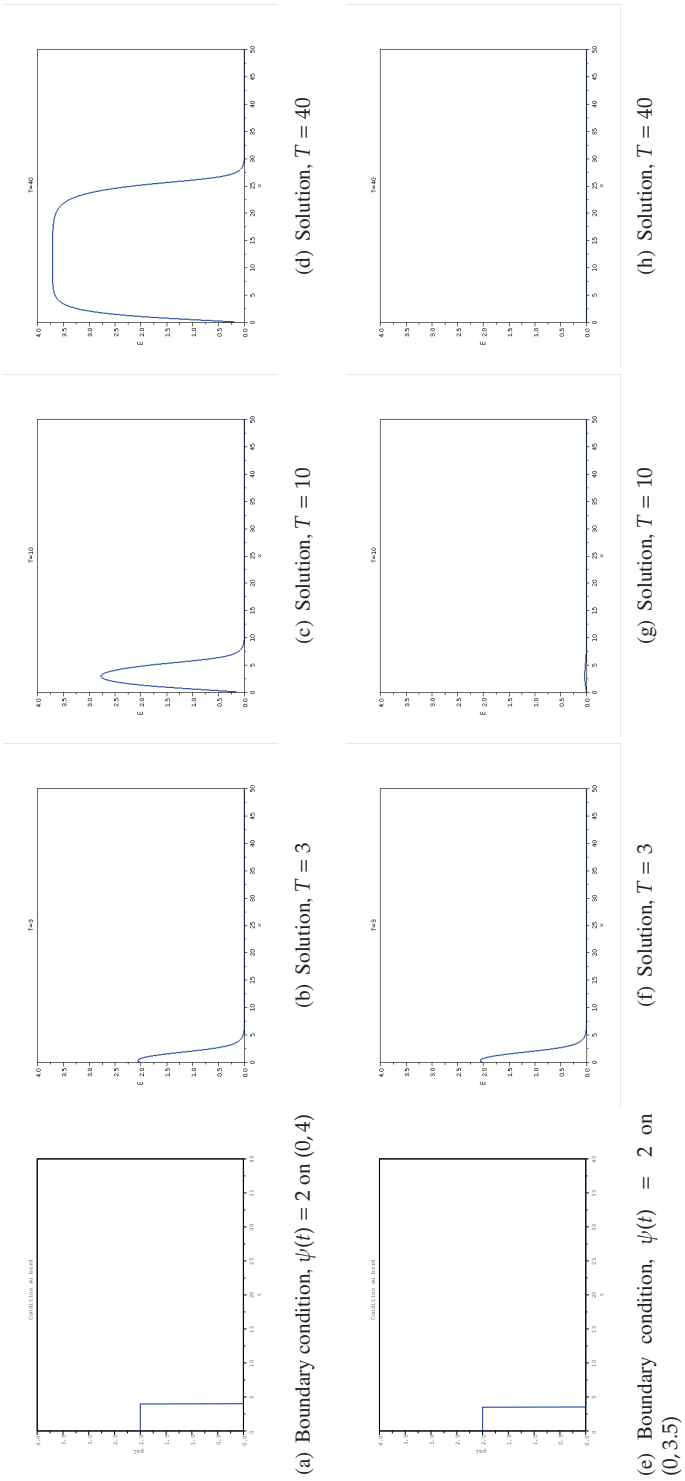


(a) Boundary condition, $\sup_t \psi(t) = 1.3$ (b) Solution, $T = 3$ (c) Solution, $T = 10$ (d) Solution, $T = 40$

(e) Boundary condition, $\sup_t \psi(t) = 1.2$ (f) Solution, $T = 3$ (g) Solution, $T = 10$ (h) Solution, $T = 40$

(i) Boundary condition, $\sup_t \psi(t) = 1.25$ (j) Solution, $T = 3$ (k) Solution, $T = 40$ (l) Solution, $T = 100$

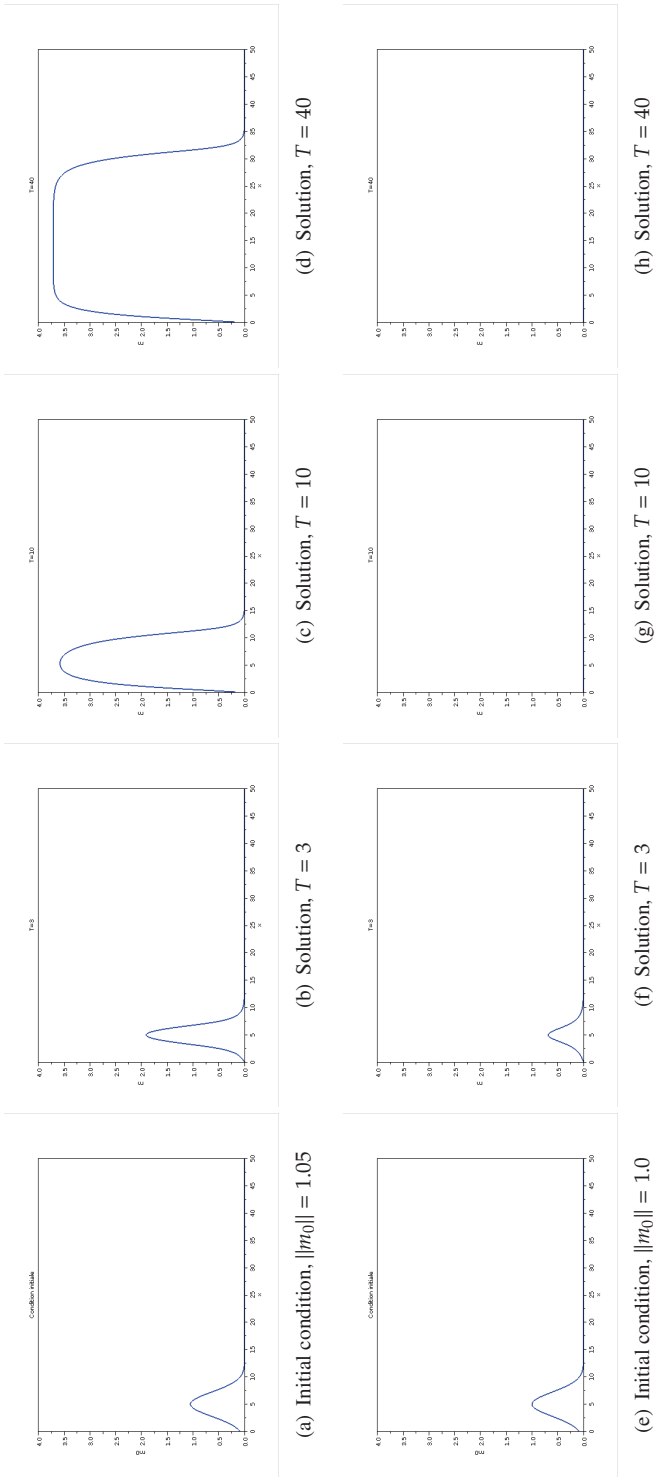
Figure 6.13: Influence of $\sup_t \psi(t)$; here $\psi(t) = \psi_{max} \dfrac{(0.5t)^4}{1 + (0.5t)^2 + (0.5t)^4}$ and $\kappa = 1.24$.



(a) Boundary condition, $\psi(t) = 2$ on $(0, 4)$ (b) Solution, $T = 3$ (c) Solution, $T = 10$ (d) Solution, $T = 40$

(e) Boundary condition, $\psi(t) = 2$ on $(0, 3.5)$ (f) Solution, $T = 3$ (g) Solution, $T = 10$ (h) Solution, $T = 40$

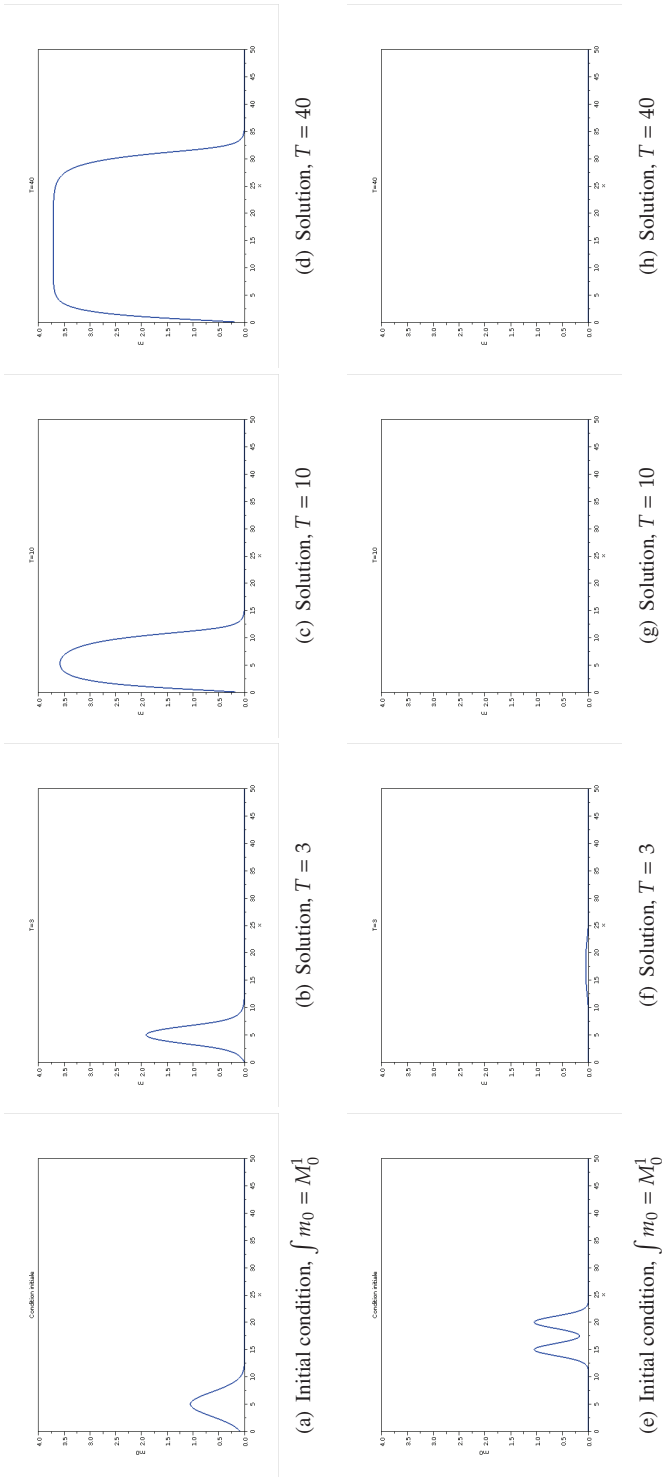
Figure 6.14: Influence of the length of the support of ψ, $\kappa = 1.24$.



(a) Initial condition, $\|m_0\| = 1.05$ (b) Solution, $T = 3$ (c) Solution, $T = 10$ (d) Solution, $T = 40$

(e) Initial condition, $\|m_0\| = 1.0$ (f) Solution, $T = 3$ (g) Solution, $T = 10$ (h) Solution, $T = 40$

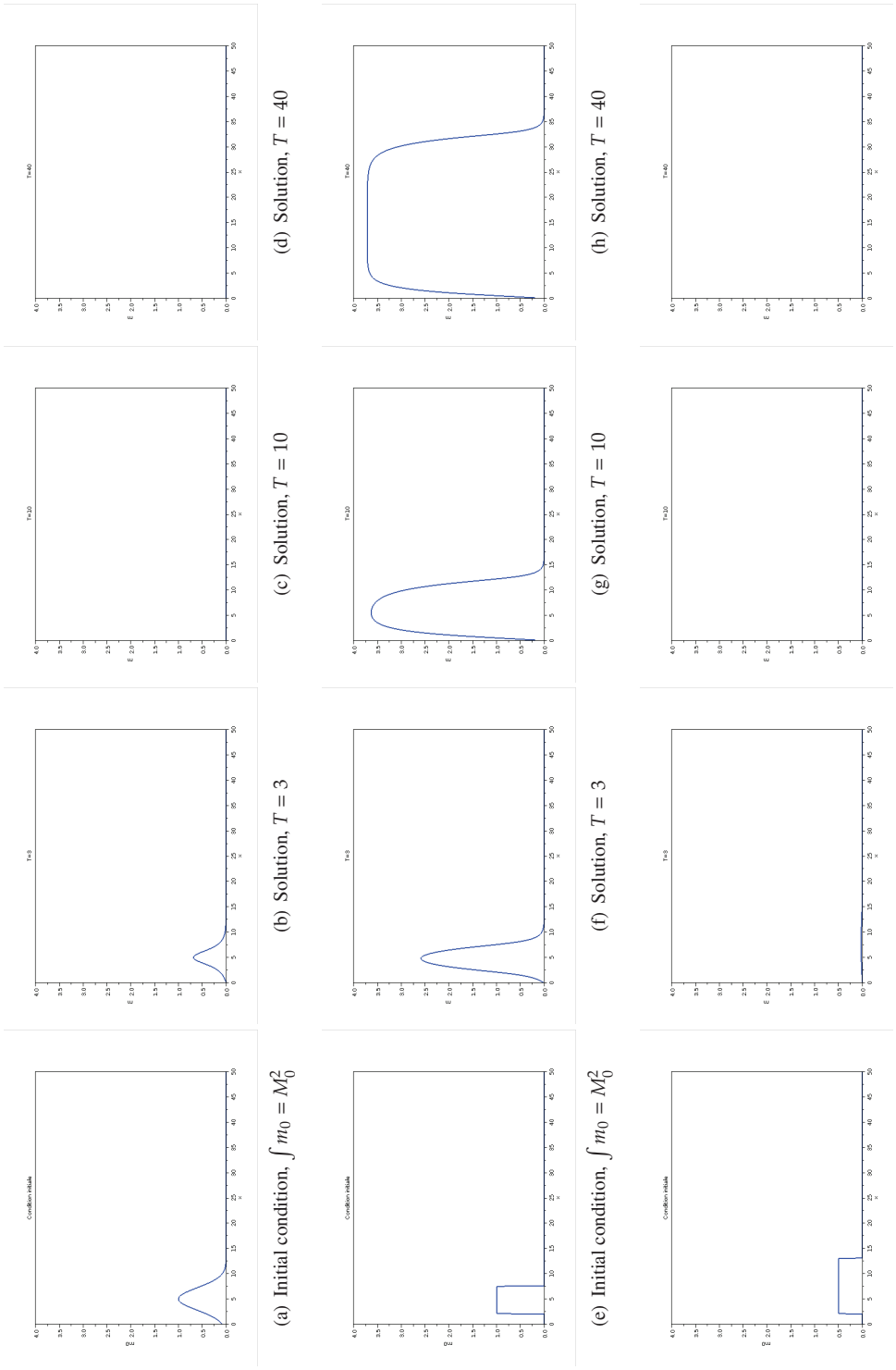
Figure 6.15: Influence of the initial condition, $\psi(t) \equiv 0$, $\kappa = 1.24$.



(a) Initial condition, $\int m_0 = M_0^1$ (b) Solution, $T = 3$ (c) Solution, $T = 10$ (d) Solution, $T = 40$

(e) Initial condition, $\int m_0 = M_0^1$ (f) Solution, $T = 3$ (g) Solution, $T = 10$ (h) Solution, $T = 40$

Figure 6.16: Influence of the shape of the initial condition for a given initial mass, $\psi(t) \equiv 0$, $\kappa = 1.24$.

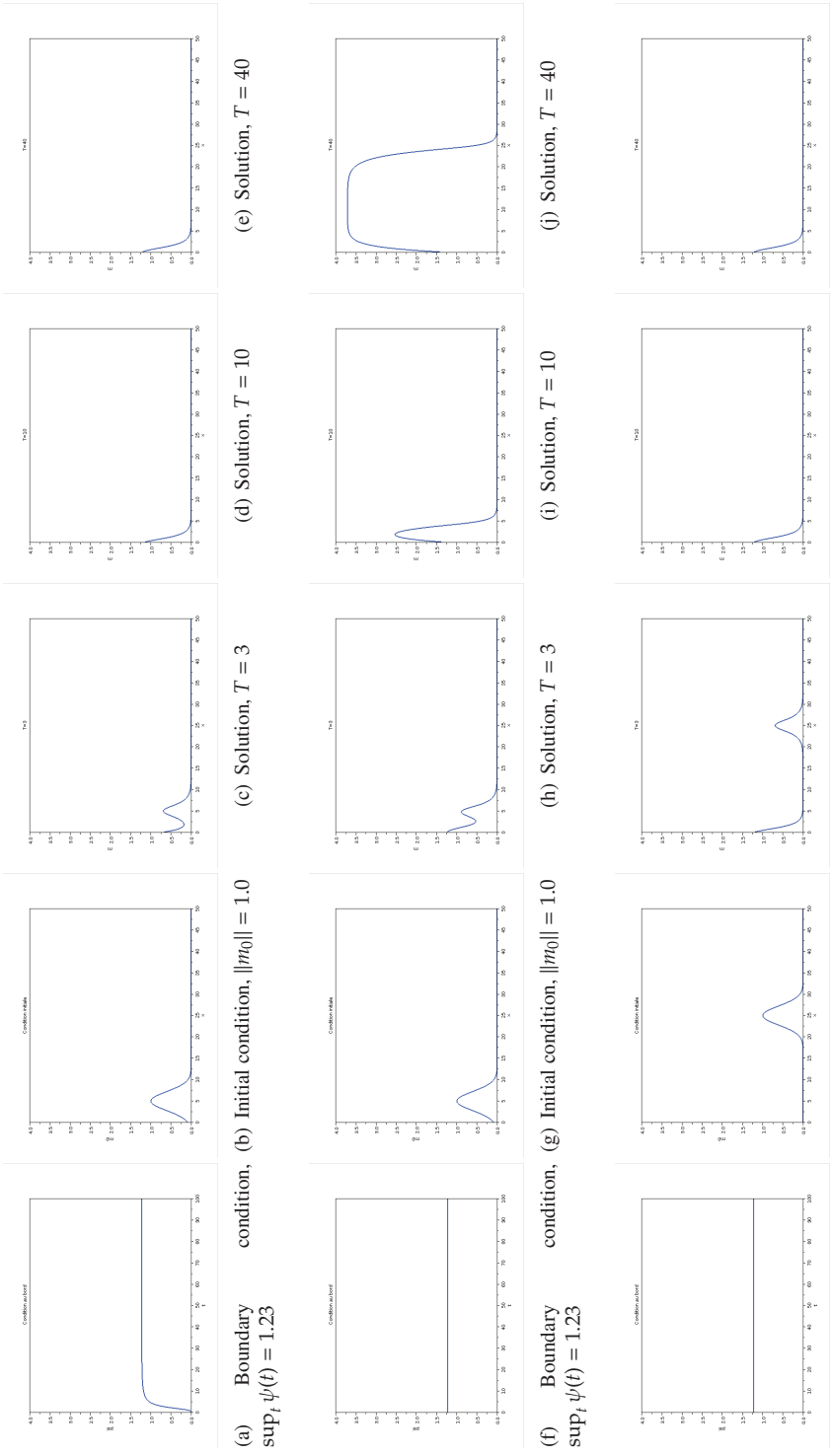


(a) Initial condition, $\int m_0 = M_0^2$ (b) Solution, $T = 3$ (c) Solution, $T = 10$ (d) Solution, $T = 40$

(e) Initial condition, $\int m_0 = M_0^2$ (f) Solution, $T = 3$ (g) Solution, $T = 10$ (h) Solution, $T = 40$

(i) Initial condition, $\int m_0 = M_0^2$ (j) Solution, $T = 3$ (k) Solution, $T = 10$ (l) Solution, $T = 40$

Figure 6.17: Influence of the shape of the initial condition for a given initial mass, $\psi(t) \equiv 0$, $\kappa = 1.24$.



(a) Boundary condition, $\sup_t \psi(t) = 1.23$ (b) Initial condition, $\|m_0\| = 1.0$ (c) Solution, $T = 3$ (d) Solution, $T = 10$ (e) Solution, $T = 40$

(f) Boundary condition, $\sup_t \psi(t) = 1.23$ (g) Initial condition, $\|m_0\| = 1.0$ (h) Solution, $T = 3$ (i) Solution, $T = 10$ (j) Solution, $T = 40$

(k) Boundary condition, $\sup_t \psi(t) = 1.23$ (l) Initial condition, $\|m_0\| = 1.0$ (m) Solution, $T = 3$ (n) Solution, $T = 10$ (o) Solution, $T = 40$

Figure 6.18: Influence of both initial condition and boundary condition. $\kappa = 1.24$.

6.5.2 Asymptotic velocity

Estimation of the asymptotic velocity

In this section we want to evaluate numerically the asymptotic velocity defined at (6.37). The method we use is a finite–difference discretization combined to Newton's method.

At first, the trajectories in the phase plane are given by Equation (6.35), but one should note that the right–hand side of (6.35) is not defined in $(0,0)$. Nevertheless, we know the slope of the trajectory at the origin (see the proof of Theorem 6.4.16 a) Step2). This is why we choose to solve numerically:

$$\begin{cases} \dfrac{\mathrm{d}p}{\mathrm{d}q} = -c^* - \dfrac{f(q)}{p} & \text{if } p \neq 0, \\ \dfrac{\mathrm{d}p}{\mathrm{d}q}(0) = \dfrac{1}{2}\left(-c^* - \sqrt{c^{*2} - 4f'(0)}\right), \\ p(0) = p(b) = 0. \end{cases} \tag{6.48}$$

We want to discretize problem (6.48). To this end we use a regular grid

$$0 = q_0 < q_1 < \cdots < q_N < q_{N+1} = b, \quad \text{with } q_{i+1} - q_i = \Delta q = \frac{b}{N+1}, \quad N \in \mathbb{N},$$

and the following forward 1st–order discretization of the problem (6.48) :

$$\begin{array}{rcll} \dfrac{p_1}{\Delta q} & = & \dfrac{\mathrm{d}p}{\mathrm{d}q}(p_0) & = \dfrac{1}{2}\left(-c - \sqrt{c^2 - 4f'(0)}\right), \\ \dfrac{p_{i+1} - p_i}{\Delta q} & = & \dfrac{\mathrm{d}p}{\mathrm{d}q}(p_i) & = -c - \dfrac{f(q_i)}{p_i} \qquad \text{for } i = 1 \ldots N-1, \\ \dfrac{-p_N}{\Delta q} & = & \dfrac{\mathrm{d}p}{\mathrm{d}q}(p_N) & = -c - \dfrac{f(q_N)}{p_N}. \end{array}$$

This leads to the following non linear system:

$$AP = B, \tag{6.49}$$

where

$$A = \frac{1}{\Delta q}\begin{pmatrix} 1 & 0 & & \ldots & & 0 \\ -1 & 1 & \ddots & & & \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & -1 & 1 \\ 0 & & \ldots & & 0 & 1 \end{pmatrix} \in M_{N+1,N}, \tag{6.50}$$

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ \vdots \\ \\ \vdots \\ p_N \end{pmatrix} \in M_{N,1} \tag{6.51}$$

and

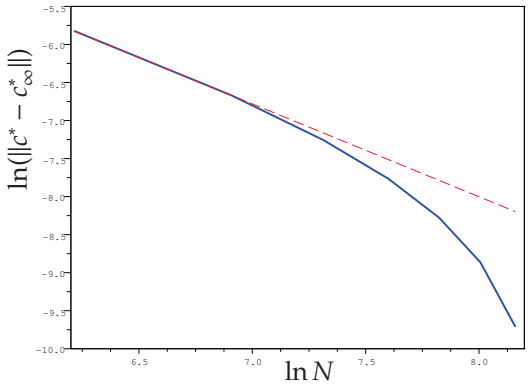


Figure 6.19: Convergence to the asymptotic velocity with respect to the size of the grid.

$$B(P,c^*) = \begin{pmatrix} \frac{1}{2}\left(-c^* - \sqrt{c^{*2} - 4f'(0)}\right) \\ -c^* - \dfrac{f(q_1)}{p_1} \\ -c^* - \dfrac{f(q_2)}{p_2} \\ \vdots \\ \vdots \\ \\ \vdots \\ -c^* - \dfrac{f(q_N)}{p_N} \end{pmatrix} \in M_{N+1,1}. \tag{6.52}$$

Note this system is not square. To compute the solution (c^*, P) of (6.48) we use a Newton procedure combined with a dichotomy procedure to initialize reasonably well. Because the finite–difference discretization we use is of order 1, at the convergence of the Newton method, which is quadratic by the way, we get an error of order Δq. To confirm this convergence rate, we plot numerically the logarithm of the error as a function of the logarithm of the size of the mesh. (See Figure. 6.19.) Using a linear regression, we find out a numerical order of convergence of $r_{num} = 1.22$.

Dependence of the asymptotic velocity on the parameter α

At Figure 6.20, we plot the asymptotic velocity as a function of α. The velocity is increasing with α, this is relevant biologically: the asymptotic velocity goes up when the degradation is reduced.

6.5.3 Influence of the Hill exponent

We want now to compare the results for several values of the Hill exponent h. Recall that until now, we have been working with $h = 4$, considered as the typical Hill exponent for the MAPK cascade. For $h = 2$ the non–linearity (see (6.15)) reads as follows:

$$z_\alpha(m) = \frac{\alpha m^2}{m^2 + 1}.$$

A Hill exponent equal to 2 appears in other biological models such as the calcium–stimulated–calcium–release mechamism which happens from calcium sites on the membrane enclosing certain fertilized amphibian eggs

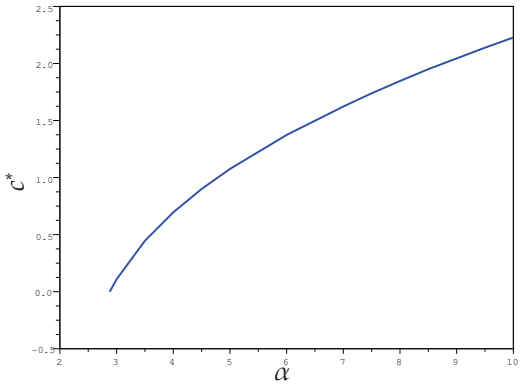


Figure 6.20: Dependence of the asymptotic velocity on the parameter α.

([Mur93]). For more details about the Hill exponent see also [GK81]. Cases $h = 2$ and $h = 4$ are both biologically relevant. To complete our simulations we choose to test the case $h = 6$ with the non–linearity defined as follows:

$$z_\alpha(m) = \frac{\alpha m^6}{m^6 + m^3 + 1}.$$

Figure 6.24 displays the shape of the functions $m \mapsto z_\alpha(m)$ for the three values of the Hill exponent $h = 2$, $h = 4$ and $h = 6$. As h increases, the slope is steeper. It would be tempting to correlate the steepness of these curves to the properties of the wave front: velocity of the wave front and speed of the change of state can be expected to vary monotonically with h (typically, the smaller h, that is when the slope is less steep, the slower the wave front would be the naive belief). We shall see that such a deduction is erroneous.

Firstly, we plot in Figure 6.21 the behavior of κ as a function of α for $h = 2$, $h = 4$ and $h = 6$. As one can see, the value of κ is lower for $h = 2$ than for $h = 4$. This means that the amount of mass needed to trigger the wave front is lower for $h = 2$ than for $h = 4$. Isolating this property is important since checking this consequence of the modeling could be accessible to experiments and the comparison would help in validating the proposed equation. The results obtained with $h = 6$ however indicate that κ is not a monotone function of h.

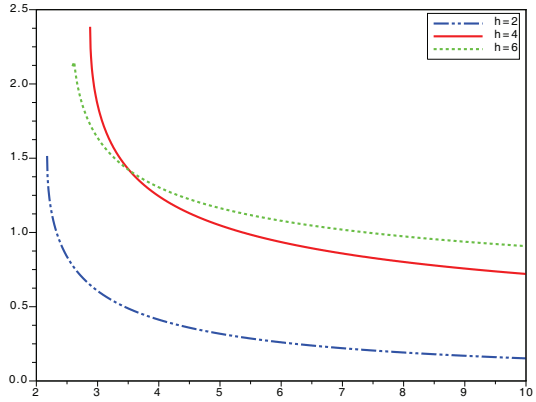


Figure 6.21: Behavior of κ (see (6.22)) with respect to α for three different values of the Hill exponent : $h = 2$, $h = 4$ and $h = 6$.

Secondly, in Figure 6.22 we plot the asymptotic velocity as a function of α for $h = 2$, $h = 4$ and $h = 6$. What could be surprising is that the asymptotic velocity c^*, which is the speed of signal propagation, is lower for $h = 4$ than for $h = 2$. The asymptotic velocity is not monotone with respect to h.

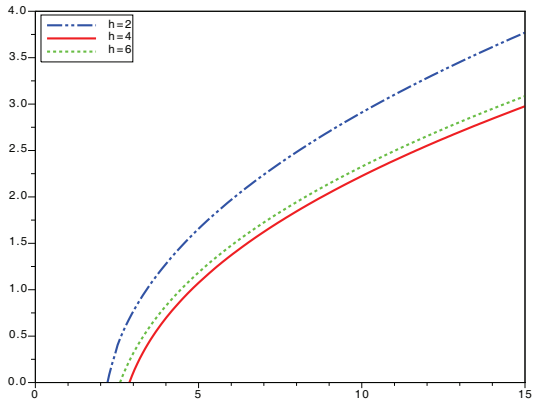


Figure 6.22: Behavior of asymptotic velocity with respect to α for three different values of the Hill exponent : $h = 2$, $h = 4$ and $h = 6$.

Thirdly, to precise the comparison for different Hill exponents we plot in Figure 6.23 the time it takes, when the asymptotic velocity is reached, to switch from the stationary state $m \equiv 0$ to $m \equiv b$. To this end we define the quantity Δ as follows:

$$\Delta = z_\alpha^{-1}(0.99 \cdot \alpha) - z_\alpha^{-1}(0.01 \cdot \alpha). \tag{6.53}$$

We plot in Figure 6.23 the quantity

$$t = \frac{\Delta}{c^*}, \tag{6.54}$$

which is expected to be proportional to the switch time. At least it is a function of h, characteristic of the state transition. Again we note that this quantity is not monotone with respect to h the Hill exponent. Nevertheless, for α small enough, the switch time is smaller for $h = 4$ than for $h = 2$.

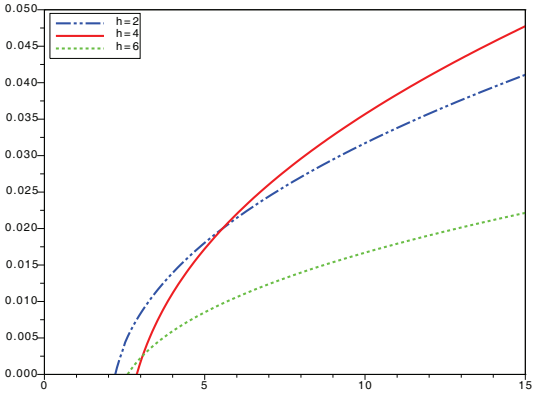


Figure 6.23: Behavior of the switch time (see (6.54)) with respect to α for three different values of the Hill exponent : $h = 2$, $h = 4$ and $h = 6$.

6.6 Conclusion and perspectives

In this work, we proposed a simplified model for signal propagation in the MAPK cascade. It reduces the problem to a reaction-diffusion equation for the kinase MAP3K, the molecule Mos in the case of *Xenopus* oocytes. The propagation of the downstream molecules of the cascade, MAP2K and MAPK, is slaved to the diffusion of Mos.

While certainly a strong simplification of the dynamics of the cascade, our model has allowed us to study a number of questions by analytic means and numerical approaches. While the model is too simplified to allow

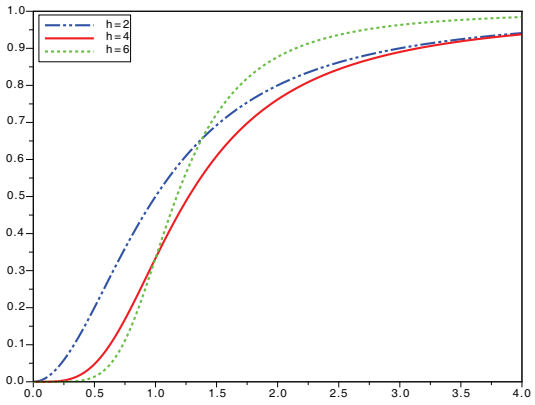


Figure 6.24: Function $m \mapsto z_\alpha(m)$ for three different values of the Hill exponent : $h = 2$, $h = 4$ and $h = 6$.

a direct comparison with experiment, its study has provided us with a number of key insights that merit further investigation in more elaborate models. Two results deserve special mention in this context. First, the role of the nonlinearity in the reaction term, hence the ultrasensitive response. The propagation speed depends on this nonlinearity, but in an non-intuitive way. There is no direct correlation between the Hill exponent h and the asymptotic speed of propagation. Definitely the speed of propagation is not monotone with respect to h (and the propagation speed is less for $h = 4$ than for $h = 2$). We do not wish to speculate too far about the possible biological consequences this fact might have, in any case it can be taken as an indication that optimization of the speed of the propagating front is not a design criterion for the signalling chain.

Second, and in our view even more intruiging, is the role played by the boundary and initial conditions. We have found that the propagation depends quite sensitively on the details of these conditions. In particular, deciding whether or not the signal propagates cannot be embodied in a single parameter (say the L^∞ or the L^1 norm of the signal).

In our view, this finding points to an important feature that has so far not been present in the discussion of the functioning of the MAPK cascade in *Xenopus* oocytes. These discussions have largely been focused around bistability indeed, a precondition for all our discussion and ultrasensitivity. Our result hints at another dimension of the signaling process : how the molecular concentration at signal input is distributed has an effect on signal propagation. Since the MAPK cascade can be initiated by progesterone in the oocyte, it might be an interesting suggestion to see whether and how spatial variations in the initial or boundary conditions might be realised and studied experimentally, and what effects they entail.

The question of the possible role of different initial or boundary conditions is of obvious biological importance. The decision of an oocyte cell to proceed into maturation may indeed be triggered by different spatial distributions of activating factors, as well as their concentrations. Two alternative scenarii are therefore possible: either the cell makes use of its sensitivity to the initial conditions and therefore can tune its reponse accordingly, or, in the opposite case, molecular architectures inside the cell should be capable of buffering concentration variations in order to control signal propagation. On the theoretical side, how the scenario delivered by our one-dimensional model will be modified in more realistic cellular geometries is, of course, an interesting open question.

6.7 Appendix A : Derivation of the expression of the reaction term $\Sigma[m(x,t)]$.

The expression of the reaction term depends on the kinetics chosen for the signaling cascade. Typically, the cascade has two phosphorylation states so that for a two-level cascade with molecules y_0 and z_0, two respective phosphorylated states exist, for y_0 these are called y_1 and y_2, and for z_0, respectively, z_1 and z_2. Recall that m denote the concentration of the molecule at the cascade entry. Two particular types of kinetics are linear and

saturated, Michaelis-Menten kinetics. In the latter case, the equations read as, following [AFS04] :

$$\dot{y}_0 = \frac{V_6 y_1}{K_6 + y_1} - \frac{V_3 m y_0}{K_3 + y_0} \tag{6.55}$$

$$\dot{y}_1 = -(\dot{y}_0 + \dot{y}_2) \tag{6.56}$$

$$\dot{y}_2 = \frac{V_4 m y_1}{K_4 + y_1} - \frac{V_5 y_2}{K_5 + y_2} \tag{6.57}$$

$$\dot{z}_0 = \frac{V_{10} z_1}{K_{10} + z_1} - \frac{V_7 y_2 z_0}{K_7 + z_0} \tag{6.58}$$

$$\dot{z}_1 = -(\dot{z}_0 + \dot{z}_2) \tag{6.59}$$

$$\dot{z}_2 = \frac{V_8 y_2 z_1}{K_8 + z_1} - \frac{V_9 z_2}{K_9 + z_2} \tag{6.60}$$

where the V_i and K_i are reaction speeds and equilibrium constants, respectively, and where the numbering of the reactions follows the scheme by [Kho00] which is also used in ref. [AFS04]. This scheme simply numbers the reactions sequentially layer by layer, first all phosphorylations, and then all dephosphorylations. Note that if equilibrium constants are much larger than the molecular concentrations, the linear regime is recovered.

As shown in [RGBB09], for a particular symmetric choice of the parameters, the system can be rendered non-dimensional. For this parameter-symmetric case, the introduction of the functions

$$w_i \equiv \frac{y_i}{1 + y_i}\,, \quad q_i \equiv \frac{z_i}{1 + z_i}\,. \tag{6.61}$$

allows to rewrite the rhs of the equations in a simple polynomial form, from which the fixed-point value conditions are easy to read off. For the variable w_i, they are given by $w_2 = m w_1 = m^2 w_0$.

Exploiting the conservation of molecules y_i as expressed by the condition $\sum_{i=0,1,2} y_i = y_T$, one finds after suitable normalization the condition

$$\frac{w_1}{1 - w_1} + \frac{w_2}{1 - w_2} + \frac{w_3}{1 - w_3} = 1\,. \tag{6.62}$$

which at the fixed-point reduces to a cubic equation, *e.g.* for w_0,

$$4m^3 w_0^3 - 3m(1 + m + m^2) w_0^2 + 2(1 + m + m^2) w_0 - 1 = 0\,. \tag{6.63}$$

which can be solved exactly. The case of linear kinetics can, however, also be recovered by ignoring the highest order terms of this equation. One then finds

$$w_0 = \frac{1}{2} \frac{1}{1 + m + m^2}\,. \tag{6.64}$$

Although the full final expression for w_0 from the cubic equation does have a more involved form, the degree of the polynomial remains two as for the linear kinetics. From the fixed-point relation, we have

$$w_2 = \frac{1}{2} \frac{m^2}{1 + m + m^2}. \tag{6.65}$$

The same calculation can be repeated for the variable q_i. One obtains an equation of the same form as (6.63) in the variable q_1, only with m replaced by

$$s = v \frac{w_2}{1 - w_2} \tag{6.66}$$

where $v \equiv V_7/V_{10}$. The same argument can then be repeated leading to

$$q_2 = \frac{1}{2}\frac{s^2}{1+s+s^2}\,. \tag{6.67}$$

Putting the equations (6.65) and (6.67) together then leads already for the approximate linear kinetics to an involved expression, however with a polynomial of maximal degree four in the variable m in both denominator and numerator. This remains the case also after the required transforms back from w_i and q_i to the original variables y_i and z_i. While the full expression obtained from the two cubic equations can still be written down analytically, it turns out to be yet more involved and less illuminating. For both types of kinetics the Hill exponent is given by the product of the number of phosphorylation levels times cascade levels, hence four. For convenience in our mathematical treatment, we replace both involved expressions by the simple expression for $\Sigma[m(x,t)]$ used in the text since we are interested chiefly in the role of ultrasensitivity, *i.e.* on the value of the Hill coefficient, on wave propagation. In order to formulate the feedback loop in which MAPK acts back on Mos, m, we follow ANGELI *et al.* [AFS04] and write

$$\dot{m} = -\gamma m + z_2 \tag{6.68}$$

in which $z_2 = \Sigma[m]$ acts as the source term.

In our derivation we have assumed a symmetry relation between equilibrium constants. This choice is judicious in order to derive the analytic formulae, however it does not influence the form of the Hill functions (*i.e.*, the order of the exponents). It is easy to see that this is indeed the case, *e.g.* by a systematic perturbation theory around the symmetric solution, see also the discussion in [Rus09].

6.8 Appendix B : Properties of q_β defined at (6.29)

Let us recall the definitions. For any $\beta \in (\kappa, b)$, where κ is defined at (6.22), the length ℓ_β is defined as follows

$$\ell_\beta = 2\int_0^\beta \frac{1}{\sqrt{2(F(\beta)-F(u))}}\,\mathrm{d}u,$$

where F is defined at (7.55). Let q_β be the solution of

$$q_\beta'' + f(q_\beta) = 0 \text{ in } \mathbb{R}^+,$$

with

$$q_\beta(0) = 0$$

and

$$\frac{1}{2}{q_\beta'}^2 + F(q_\beta) = F(\beta). \tag{6.69}$$

Let us now show that these conditions imply

$$q_\beta > 0,\ q_\beta(0) = q_\beta(\ell_\beta) = 0 \text{ and } q_\beta(x) \le q_\beta\left(\frac{\ell_\beta}{2}\right) = \beta \text{ on } (0,\ell_\beta)$$

$\underline{q_\beta\left(\frac{\ell_\beta}{2}\right) = \beta}$: (a) First, we determine the zeros of the first derivative of q_β, q_β'. Thanks to the first integral (6.69), if there exists x_0 such that $q_\beta'(x_0) = 0$ then necessarily $F(q_\beta(x)) = F(\beta)$. Since F is a bijection from $[\kappa, b]$ to $[0, F(b)]$ and non positive on $[0,\kappa]$ (see Figure 6.25), this implies $q_\beta(x_0) = \beta$.

(b) From the definition,

$$\frac{\ell_\beta}{2} = \int_0^\beta \frac{1}{\sqrt{2(F(q_\beta(x_0))-F(u))}}\,\mathrm{d}u.$$

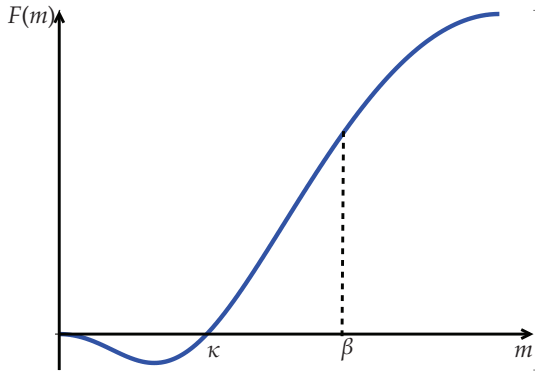


Figure 6.25: Function $F(m) = \int_0^x f(u)\,\mathrm{d}u$.

Since $q'_\beta > 0$ if $q_\beta \in (0,\beta)$, we can set the change of variable $u = q_\beta(y)$. On $(0,\beta)$, we have $q'_\beta(y) = \sqrt{2(F(q_\beta(x_0)) - F(q_\beta(y)))}$, thus we obtain

$$\frac{\ell_\beta}{2} = \int_0^{x_0} \mathrm{d}y = x_0.$$

Finally, this shows that

$$q_\beta\left(\frac{\ell_\beta}{2}\right) = \beta.$$

$\underline{q_\beta(\ell_\beta) = 0}$: We showed that $q'_\beta > 0$ on $\left(0, \frac{\ell_\beta}{2}\right)$, and q'_β vanishes and changes of sign at $\frac{\ell_\beta}{2}$. In fact $q_\beta(x) \le \beta$ for $x \ge \frac{\ell_\beta}{2}$ and we proved previously that q' vanishes in x_0 if and only if $q_\beta(x_0) = \beta$. Finally, $q'_\beta < 0$ on $\left(\frac{\ell_\beta}{2}, \beta\right)$. Let us define now $r(u) = q_\beta\left(\frac{\ell_\beta}{2} - u\right)$ and $s(u) = q_\beta\left(\frac{\ell_\beta}{2} + u\right)$. They are both solution of the following Cauchy problem:

$$\begin{aligned} y'' + f(y) &= 0 \\ y(0) &= \beta, \\ y'(0) &= 0. \end{aligned}$$

As a consequence, $r(u) = s(u)$ on $\left[0, \frac{\ell_\beta}{2}\right]$. Thus,

$$\begin{aligned} q_\beta(\ell_\beta) &= \int_0^{\ell_\beta} q'_\beta(t)\,\mathrm{d}t \\ &= \int_0^{\frac{\ell_\beta}{2}} q'_\beta(t)\,\mathrm{d}t + \int_{\frac{\ell_\beta}{2}}^{\ell_beta} q'_\beta(t)\,\mathrm{d}t, \\ &= \int_0^{\frac{\ell_\beta}{2}} q'_\beta\left(\frac{\ell_\beta}{2} - u\right)\mathrm{d}u + \int_0^{\frac{\ell_\beta}{2}} q'_\beta\left(\frac{\ell_\beta}{2} + u\right)\mathrm{d}u \\ &= -\int_0^{\frac{\ell_\beta}{2}} r'(u)\,\mathrm{d}u + \int_0^{\frac{\ell_\beta}{2}} s'(u)\,\mathrm{d}u \\ &= 0. \end{aligned}$$

Acknowledgments

My co-workers and I want to thank Frédéric Lagoutière for his fruitful remarks.

Bibliography

[AFS04] David Angeli, James E. Ferrell, and Eduardo D. Sontag.
Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems.
Proceedings of the National Academy of Sciences of the United States of America, 101(7):1822–1827, 2004.

[AW75] Donald G. Aronson and Hans F. Weinberger.
Nonlinear diffusion in population genetics, combustion, and nerve propagation.
Lecture Notes in Mathematics, 446:5–49, 1975.

[BS09] Matthew D. Brown and David B. Sacks.
Protein scaffolds in map kinase signalling.
Cellular Signalling, 21(4):462 – 469, 2009.

[FJPK+09] James E. Ferrell Jr., Joseph R. Pomerening, Sun Young Kim, Nikki B. Trunnell, Wen Xiong, Chi-Ying Frederick Huang, and Eric M. Machleder.
Simple, realistic models of complex biological processes: Positive feedback and bistability in a cell fate switch and a cell cycle oscillator.
FEBS Letters, 583(24):3999 – 4005, 2009.
Systems Biology - Nobel Symposium 146.

[GK81] Albert Goldbeter and Daniel E. Koshland.
An amplified sensitivity arising from covalent modification in biological systems.
Proceedings of the National Academy of Sciences of the United States of America, 78(11):6840–6844, 1981.

[KHK10] Boris N. Kholodenko, John F. Hancock, and Walter Kolch.
Signalling ballet in space and time.
Nature Reviews Molecular Cell Biology, 11:414–426, 2010.

[Kho00] Boris N. Kholodenko.
Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades.
Eur J Biochem, 267(6):1583–1588, Mar 2000.

[Kho02] Boris N. Kholodenko.
Map kinase cascade signaling and endocytic trafficking: a marriage of convenience?
April 2002.

[Kho06] Boris N. Kholodenko.
Cell-signalling dynamics in time and space.
Nat Rev Mol Cell Biol, 7:165–176, March 2006.

[MTHK06] Nick I. Markevich, Mikhail A. Tsyganov, Jan B. Hoek, and Boris N. Kholodenko.
Long-range signaling by phosphoprotein waves arising from bistability in protein kinase cascade.
Molecular Systems Biology, 2006.
Article number : 61.

[Mur93] James D. Murray.

Mathematical Biology, 2d ed., volume 19 of *Biomathematics*.
Springer, New York, 1993.

[RGBB09] Christophe Russo, Vasile Claudiu Giuraniuc, Ralf Blossey, and Jean-François Bodart.
On the equilibria of the mapk cascade: Cooperativity, modularity and bistability.
Physica A: Statistical Mechanics and its Applications, 388(24):5070 – 5080, 2009.

[Rus09] Christophe Russo.
Dynamique et modularité de la voie Mitogen Activated Protein Kinase dans les ovocytes de Xénope : modélisation et approches expérimentales.
PhD thesis, Université de Lille 1, U.F.R. de Biologie, 2009.

Quatrième partie

Modèle de Patlak–Keller–Segel 1D

Chapitre 7

Description et simulation du système de Patlak–Keller–Segel modifié 1D après « blow–up »

Dans ce chapitre, nous nous intéressons à l'équation de Patlak–Keller–Segel en dimension $d = 1$ où la concentration chimique est donnée par la convolution avec le noyau logarithmique. Les mêmes équations, posées non pas sur $\mathbb{R}$ mais $\mathbb{R}^2$, correspondent au cas classique originellement décrit par KELLER et SEGEL. Ce travail se situe à la croisée de deux travaux antérieurs. L'un, réalisé par DOLBEAULT *et al.* [DS09], décrit le système de Keller–Segel en dimension $d = 2$ après blow–up à l'aide de la notion de mesure de défaut développée par POUPAUD. L'autre, effectué par BLANCHET *et al.* [BCC08], donne un schéma numérique pour le système étudié ici valable avant explosion et fondé sur une formulation du problème en termes de flots de gradient pour la métrique de Wasserstein. Dans ce qui suit, nous adapterons les idées de DOLBEAULT *et al.* au cas de la dimension $d = 1$ pour obtenir des équations décrivant le système après explosion, puis nous utiliserons ce résultat pour étendre le schéma de BLANCHET *et al.* aux temps après explosion.
Ce travail a été obtenu collaboration avec VINCENT CALVEZ et JOSÉ ANTONIO CARRILLO.

7.1 Introduction

Le travail présenté dans ce chapitre touche aux phénomènes de chimiotactisme. Le chimiotactisme est la tendance des cellules ou des organismes mobiles à se déplacer dans une direction déterminée sous l'influence de substances chimiques, appelées chimioattracteurs. Ce phénomène est essentiel en biologie : trouver de la nourriture (glucose) en se dirigeant vers sa concentration la plus élevée par exemple. Le chimiotactisme joue également un rôle clé dans l'organisation spatiale des cellules, voire dans l'auto–organisation lorsque le chimioattracteur est émis par les cellules elles–mêmes (voir par exemple le cas du *Dictyostelium discoideum*[Wei04, Wei09], Figure 7.1).

En biologie humaine, nous pouvons citer l'exemple de l'angiogenèse[8] lors de la croissance tumorale. Lors des premiers stades de son développement, la tumeur se nourrit en puisant les nutriments nécessaires à sa croissance dans son environnement par diffusion à partir des vaisseaux sanguins existants. Cette phase de croissance est dite avasculaire. Cependant, l'oxygène et les nutriments ne pouvant diffuser à plus de quelques millimètres des vaisseaux sanguins, les ressources nécessaires à la poursuite du développement tumoral deviennent rares. La tumeur déclenche alors le processus angiogénique en émettant des facteurs de croissance de l'endothélium vasculaire (VEGF) qui agissent en attirant les vaisseaux sanguins issus du réseau vasculaire existant vers la tumeur, afin de

[8] Processus de formation de nouveaux vaisseaux sanguins à partir d'un réseau vasculaire préexistant.

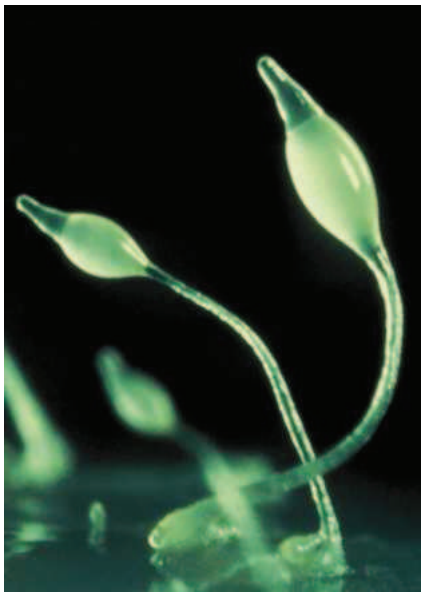


FIG. 7.1: *Dictyostelium discoideum*. Agrégats formés d'environ 10^5 organismes unicellulaires. Ce"corps pluricellulaire" est constitué principalement de types de cellules. Les cellules du domaine "spore" forment le pied du champignon et sont sacrifiées dans ce processus (approx. 80% des cellules). Les cellules du domaine "stalk" au sommet du champignon sont susceptibles de coloniser d'autres endroits (approx. 20% des cellules).

l'irriguer et lui apporter les nutriments nécessaires à sa croissance [Bil09, BP00]. Toujours dans le domaine de la biologie humaine, nous pouvons également citer l'exemple du développement du réseau de connexions neuronales.

Dans ce qui suit, nous nous sommes intéressés plus particulièrement à un modèle d'auto–organisation, à savoir une population de cellules qui produit elle–même le signal chimique qui les attire.

7.1.1 Le modèle parabolique de Patlak– Keller–Segel

Nous considérons une population de cellules en déplacement, qui interagissent entre–elles via un signal chimique en remontant le gradient de potentiel. Nous supposerons de plus une diffusion linéaire des cellules et des molécules du signal chimique. Le système couplé obtenu, dit modèle de Patlak–Keller–Segel (PKS), est le suivant :

$$\begin{cases} \dfrac{\partial \rho}{\partial t} &= \Delta \rho - \nabla \cdot (\chi \rho \nabla c), \quad t > 0,\ x \in \Omega \subset \mathbb{R}^2, \\ \varepsilon \dfrac{\partial c}{\partial t} &= \Delta c + \rho - \alpha c, \end{cases} \tag{7.1}$$

où ρ représente la densité de cellules et c la concentration du signal chimique. Dans le cas où le domaine Ω est borné et régulier, le système doit être complété par des conditions aux bords. Dans ce cas, on choisit des conditions aux bords de Neumann homogènes pour la densité de cellules, permettant de garantir la conservation de la masse totale, et de type Neumann ou Dirichlet pour la concentration de l'espèce chimique. Dans le système (7.1), le paramètre $\chi > 0$ représente la chimiosensibilité, ε est le rapport des coefficients de diffusion pour les cellules avec le coefficient de diffusion de l'espèce chimique dans le milieu : $\varepsilon = D_\rho / D_c$. IL est petit dans le cas des cellules de grande taille qui ont un mouvement lent (mouvement amiboide ou mesenchymal). En revanche ε est d'ordre 1 pour des bactéries en milieu liquide. Enfin $\alpha \geq 0$ représente le taux de dégradation. Le dernier paramètre important est la masse totale du système $M = \displaystyle\int_\Omega \rho(t=0,x)\,\mathrm{d}x$ qui est conservée au cours du temps.

Dans la suite nous prendrons pour Ω tout entier. De plus, nous nous placerons dans le cas quasi–stationnaire [JL92], qui modélise le fait qu'expérimentalement le coefficient de diffusion pour le chimioattractant est très grand,

i.e. ε est très petit. Nous fixons donc maintenant $\varepsilon = 0$. Nous supposerons dans la suite que la dégradation de l'espèce chimique est négligeable et ainsi $\alpha = 0$.

Dans ce cas ($\varepsilon = \alpha = 0$ et $\Omega = \mathbb{R}^2$), c est définie par la convolution avec la solution fondamentale de l'équation de Poisson,

$$c(t,x) = -\frac{1}{2\pi}\int_{\mathbb{R}^2} \log|x-y|\,\rho(t,y)\,\mathrm{d}y. \tag{7.2}$$

Ce modèle rend compte d'un phénomène de masse critique observé en réalité. Intuitivement, deux phénomènes sont mis en balance : la diffusion des cellules par agitation brownienne d'une part, et l'attraction des cellules les unes aux autres via le potentiel du signal chimique d'autre part. Nous devinons que si les cellules sont assez nombreuses elles vont reconnaître leurs présences respectives et s'attirer pour former des agrégats. Autrement dit, nous suspectons qu'au delà d'une masse critique une organisation spatialement non homogène de l'espace par les cellules émerge. La partie non–linéaire de l'équation (7.1) exhibe une boucle de rétro–action positive : là où il y a beaucoup de cellules, beaucoup de chimioattractant est produit attirant de plus en plus de cellules, etc. Ce processus est contre–balancé par la diffusion, qui tend quant à elle à disperser les cellules. Pour se faire une idée simple de cette compétition entre la dispersion par diffusion et l'agrégation par chimiotactisme, nous pouvons faire l'analogie avec un système dynamique très simple qui combine un terme linéaire négatif (le Laplacien) avec un terme quadratique positif (l'interaction). Par exemple $\dot{x} = -x + x^2$ présente une dichotomie très simple : si $x_0 < 1$ il y a existence globale des solutions, et convergence vers zéro. En revanche si $x_0 > 1$ il y a explosion en temps fini du système.

Dans le cas du système de Patlak–Keller–Segel, le phénomène de masse critique est précisé d'un point de vue mathématique par le théorème suivant [BDP06, DP04] :

Théorème 7.1.1 (Dobeault-Perthame '04) *Dans l'espace entier, $\Omega = \mathbb{R}^2$ avec $\alpha = 0$ et $\varepsilon = 0$, pour une donnée initiale $\rho_0\left(|\log\rho_0| + (1+|x|^2)\right) \in L^1$, nous avons l'alternative suivante : la solution est globale en temps si $\chi M < 8\pi$, alors qu'elle explose en temps si $\chi M > 8\pi$.*

7.1.2 Le modèle de Patlak–Keller–Segel modifié

Pour le modèle (7.1), le comportement du système dépend beaucoup de la dimension. En effet, en dimension $d = 1$, l'explosion ne peut avoir lieu [HP04] tandis qu'en dimension $d = 3$, l'explosion peut arriver pour des masses arbitrairement petites [CPZ04].

Dans [CPSt], les auteurs suggèrent une variante du modèle de Patlak–Keller–Segel pour lequel la concentration chimique est donnée par la convolution avec le noyau logarithmique,

$$\begin{cases} \dfrac{\partial\rho}{\partial t} &= \Delta\rho - \nabla\cdot(\chi\rho\nabla c), \quad t>0, x\in\Omega\subset\mathbb{R}^d, \\ c &= -\dfrac{1}{d\pi}\log|z| * \rho. \end{cases} \tag{7.3}$$

En particulier, le gradient de concentration ∇c en dimension 1 est donné par l'opposé de la transformée de Hilbert de ρ, $\nabla c = -\mathcal{H}\rho$. Remarquons aussi, que dans le cas de la dimension $d = 2$, ce système correspond à (7.1) via la formulation (7.2) à l'aide du noyau de Poisson. Pour ce système modifié, les auteurs montrent un résultat de masse critique unifié analogue à celui donné au Théorème 7.1.1 dans le cas 2D "classique".

Théorème 7.1.2 (Calvez–Perthame–Sharifi tabar '07) *Si Ω est l'espace tout entier $\mathbb{R}^d$ et la donnée initiale vérifie $\rho_0\left(|\log\rho_0| + (1+|x|^2)\right) \in L^1$, alors la solution de* (7.3) *est globale si $\chi M < 2d^2\pi$, et il y a explosion en temps fini si $\chi M > 2d^2\pi$.*

Ces systèmes, posés en dimensions différentes, ont l'avantage de partager, à une constante près, la même énergie libre, ce qui permet notamment d'obtenir ce résultat uniforme. C'est ce système posé en dimension $d = 1$ qui fera l'objet du travail qui suit. Numériquement, il est intéressant d'avoir une version unidimensionnelle du système de Patlak–Keller–Segel partageant la même énergie libre. En effet les calculs 1D étant beaucoup moins coûteux, il est vraiment possible de réaliser des études numériques fines du système. De plus, cette formulation a permis d'explorer de nouvelles pistes pour traiter les fonctionnelles d'énergie libre et les flots de gradients associés

pour la distance de Wasserstein [BCC08]. Ce travail développe un schéma numérique 1D valable jusqu'au temps de l'explosion.

L'objectif de ce travail est en s'inspirant des travaux [DS09], de décrire le système de Patlak–Keller–Segel après explosion par passage à la limite dans un système régularisé, passage à la limite faisant intervenir la notion de mesure de défaut. Nous voulons ensuite adapter les techniques de [BCC08] au problème obtenu, afin de simuler le système après « blow–up ». Mentionnons ici les travaux de VELÁZQUEZ sur l'étude du système de Keller–Segel après explosion [Vel04c, Vel04b, Vel04a].

Le chapitre est organisé comme suit. Dans une première partie, nous détaillerons la notion de mesure de défaut décrite notamment par POUPAUD [Pou02]. Cette notion a été utilisée afin de décrire le système de Patlak–Keller–Segel après « blow–up » dans le cas de la dimension $d = 2$ [DS09]. Nous adapterons ce travail au cas de la dimension $d = 1$. Dans une deuxième partie, nous décrirons le schéma de Jordan–Kinderlehrer–Otto [Ott01, JKO98], qui utilise les notions de flots de gradients associés pour la distance de Wasserstein, ce qui nous permettra de décrire précisément le schéma développé dans [BCC08]. À partir de ces deux outils, nous introduirons alors un schéma pour approcher numériquement les solutions du système unidimensionnel après explosion.

Remarque 14 *Dans toute la suite et sans perte de généralité, nous supposerons*

$$M = \int_{\mathbb{R}} \rho_0(t = 0, x)\,\mathrm{d}x = 1. \tag{7.4}$$

Ainsi le comportement du système ne dépend que de χ.

7.2 Description du système après « blow–up »

Comme nous l'avons mentionné à la section précédente, pour le système de Patlak–Keller–Segel modifié (7.3) en dimension $d = 1$, il y a explosion en temps fini des solutions si $\chi > 2\pi$. L'explosion est un phénomène de concentration, où l'on voit apparaître des agrégats. Dans ce qui suit nous montrons l'existence globale de solutions généralisées, autorisant les solutions mesures. Cela étend en quelque sorte le concept de solution aux temps après explosion. Le résultat d'existence est une application de la théorie développée par POUPAUD [Pou02]. Ce travail a déjà été fait pour le système de Patlak–Keller–Segel 2D par DOLBEAULT et SCHMEISER [DS09]. Pour ce résultat d'existence, la densité de cellules est caractérisée par l'ajout d'une mesure de défaut, qui s'annule dans le cas de distributions régulières. Ces solutions globales sont construites comme limite de solutions d'un problème régularisé.

Partant de ce résultat, une formulation forte est dérivée sous l'hypothèse que la solution généralisée s'écrit comme somme d'une fonction continue et d'un certain nombre d'agrégats, modélisés par des masses de Dirac se mouvant le long de chemin réguliers. Nous verrons que contrairement à ce qui est fait dans [DS09], nous ne pouvons pas supposer la partie continue comme régulière car dans ce cas le système obtenu ne vérifie plus l'équation de conservation de la masse. La bonne hypothèse sera de supposer la partie continue, C^1 par morceaux, avec les discontinuités situées aux points de concentration. Elle devra en plus s'annuler en ces points.

7.2.1 Mesures de défaut

En dimension $d = 2$, d'un point de vue biologique, les points de concentrations représentent des agrégats de cellules, et la description de la dynamique de ces agrégats et de leurs interactions à la fois entre agrégats et avec les cellules non concentrées poussent DOLBEAULT et SCHMEISER [DS09] à étudier des modèles régularisés, qui parfois donnent des descriptions plus précises de la dynamique des processus biologiques. Nous suivons ici leur raisonnement et appliquons leurs travaux au cas 1D.

Pour cela, dans cette partie, nous nous intéressons à l'étude du problème régularisé :

$$\partial_t \rho^\varepsilon + \partial_x \cdot (\chi \rho^\varepsilon \, \partial_x S_\varepsilon[\rho^\varepsilon] - \partial_x \rho^\varepsilon) = 0, \tag{7.5}$$

$$S_\varepsilon[\rho^\varepsilon](t, x) = -\frac{1}{\pi} \int_{\mathbb{R}} \log(|x - y| + \varepsilon) \rho^\varepsilon(t, y)\,\mathrm{d}y, \tag{7.6}$$

complété par la condition initiale

$$\rho^\varepsilon(t=0) = \rho_0 \in L^1_+(\mathbb{R}) \cap L^\infty(\mathbb{R}). \tag{7.7}$$

La régularisation consiste en l'ajout d'une petite quantité ε dans le logarithme afin de s'éloigner de la discontinuité en 0.

Notre objectif dans cette section est de caractériser la limite des solutions à ce problème lorsque ε tend vers zéro, globalement en temps et quelque soit la valeur de χ.

7.2.2 Convergence de mesures : rappels et notations

Dans cette courte partie, nous rappelons quelques définitions de convergences dans le cadre mesure et nous rappelons le théorème de Prokhorov qui permet d'obtenir de la compacité dans un espace de mesures.

Définition 7.2.1 *Soit $\mathcal{M}^1(\mathbb{R}^d)$ l'ensemble des mesures bornées sur $\mathbb{R}^d$, à savoir l'ensemble des formes linéaires continues sur $C^0(\mathbb{R}^d)$. Nous notons $\mathcal{M}^1_+$ le sous–ensemble des mesures bornées positives de $\mathbb{R}^d$.*

Soit μ^n une suite de mesures de $\mathcal{M}^1(\mathbb{R}^d)$.

Définition 7.2.2 *On dit que $\mu^n \in \left(\mathcal{M}^1(\mathbb{R}^d)\right)^{\mathbb{N}}$* converge vaguement *vers $\mu \in \mathcal{M}^1(\mathbb{R}^d)$, si elle converge dans $\mathcal{M}^1(\mathbb{R}^d) - faible - \star$, à savoir, pour toute fonction $\varphi \in C^0_c(\mathbb{R}^d)$*

$$\int_{\mathbb{R}^d} \varphi(x)\mu^n(x)\,\mathrm{d}x \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} \varphi(x)\mu(x)\,\mathrm{d}x \tag{7.8}$$

Pour des suites de mesures positives $\mu^n \in \mathcal{M}^1_+(\mathbb{R}^d)$ nous disposons d'une notion de convergence plus précise :

Définition 7.2.3 *On dit que $\mu^n \in \mathcal{M}^1_+(\mathbb{R}^d)$* converge étroitement *vers $\mu \in \mathcal{M}^1_+(\mathbb{R}^d)$ si elle converge vaguement et si de plus*

$$\lim_{n\to\infty} \mu^n(\mathbb{R}^d) \to \mu(\mathbb{R}^d).$$

Dans ce cas, (7.8) *est vérifiée pour toute fonction test φ continue bornée.*

Définition 7.2.4 *On dit qu'une suite $\mu^n \in \left(\mathcal{M}^1_+(\mathbb{R}^d)\right)^{\mathbb{N}}$ est* tendue *('tightly bounded' en anglais) s'il existe une constante $M > 0$ telle que*

$$\begin{aligned} &\forall n \geq 1,\ \mu^n(\mathbb{R}^d) \leq M, \\ &\sup_{n\in\mathbb{N}} \mu^n(\mathbb{R}^d \backslash B_R) \underset{R\to\infty}{\longrightarrow} 0, \end{aligned} \tag{7.9}$$

où B_R représente la boule de rayon R dans $\mathbb{R}^d$, centrée à l'origine.

La condition (7.9) est le critère de Prokhorov qui assure que la suite est compacte pour la topologie étroite sur les mesures.

Théorème 7.2.1 (Prokhorov) *Soit $(\mu^n)_{n\geq 1} \in \left(\mathcal{M}^1_+(\mathbb{R}^d)\right)^{\mathbb{N}}$ une suite de mesures tendue. Il existe une sous–suite (μ^{n_k}) et une mesure $\mu \in \mathcal{M}^1_+(\mathbb{R}^d)$, telles que (μ^{n_k}) converge étroitement vers μ.*

En conséquence, une suite de mesures positives qui converge vaguement est étroitement convergente si et seulement si elle est tendue.

Afin de préciser la problématique qui va nous amener à introduire la notion de mesure de défaut nous citons le résultat suivant :

Lemme 7.2.5 *Soit $(\mu^n)_{n\geq 1} \in \left(\mathcal{M}^1_+(\mathbb{R}^d)\right)^{\mathbb{N}}$ une suite de mesures qui converge vaguement vers μ. Soit $F \subset \mathbb{R}^d$ tel que $\mu(F) = 0$. Alors pour toute fonction φ Borélienne de $\mathbb{R}^d$ à support compact et continue en tout point de $\mathbb{R}^d \backslash F$, on a*

$$\int \varphi(x)\mu^n(x)\,\mathrm{d}x \underset{m\to\infty}{\longrightarrow} \int \varphi(x)\mu(x)\,\mathrm{d}x.$$

Le lecteur intéressé trouvera une preuve de ce Lemme dans [Sch81], chap IV, paragraphe 6.

L'idée sous-jacente au travail qui suit est la suivante [Pou02]. Supposons que $\mu(F) \neq 0$ et que les discontinuités des fonctions φ sur F soient précisées. Le Lemme 7.2.5 ne peut être appliqué et les limites des suites $\int \varphi(x)\mu^n(x)\,\mathrm{d}x$ doivent être caractérisées à l'aide d'une mesure de défaut.

7.2.3 Estimations *a priori* et mesure de défaut diagonale

Théorème 7.2.2 *Pour tout $\varepsilon > 0$, le problème* (7.5)–(7.7) *possède une solution globale vérifiant pour tout $t > 0$*

$$\rho^\varepsilon(\cdot,t) \in L^1_+(\mathbb{R}) \cap L^\infty(\mathbb{R}), \tag{7.10}$$

$$\|\rho^\varepsilon(\cdot,t)\|_{L^1(\mathbb{R})} = \|\rho_0\|_{L^1(\mathbb{R})} = M, \tag{7.11}$$

et

$$\|\rho^\varepsilon(\cdot,t)\|_{L^\infty(\mathbb{R})} \leq c\left(1 + \frac{1}{\varepsilon^2}\right) \tag{7.12}$$

où c est une constante indépendante de ε.

Preuve. La régularisation du potentiel permet de raisonner, au moins sur un intervalle de temps assez petit, par des arguments de point fixe standards. Nous obtenons ainsi une solution régulière. Nous allons montrer que cette solution vérifie des estimations uniformes, qui permettent d'établir que la solution est définie pour tout temps positif. La conservation de la masse est une conséquence immédiate de la forme conservative de l'équation et de la régularité de la solution.
Grâce à la conservation de la masse (7.11), nous obtenons :

$$|\,\partial_x S_\varepsilon[\rho^\varepsilon](t,x)| = \left|-\frac{1}{\pi}\int_{\mathbb{R}} \frac{\rho^\varepsilon(t,y)\,\mathrm{d}y}{|x-y|+\varepsilon} \times \frac{x-y}{|x-y|}\right| \leq \frac{M}{\pi\varepsilon}.$$

Nous montrons alors (7.12) à l'aide du lemme suivant (Lemme 1 dans [HPS07]).

Lemme 7.2.6 (Hillen–Painter–Schmeiser'07) *Soient un champ de vecteurs $v : \mathbb{R}^n \times (0,\infty) \to \mathbb{R}^n$ uniformément borné et $u_0 \in L^1_+(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. Alors la solution du problème :*

$$\begin{cases} u_t = \nabla \cdot (\nabla u - uv), \\ u(t=0) = u_0, \end{cases}$$

vérifie $u \in L^\infty((0,\infty)\times\mathbb{R}^n)$ et

$$\sup_t \|u\|_\infty \leq C\left(\|u_0\|_1, \|u_0\|_\infty, \sup_t \|v\|_\infty, n\right).$$

L'existence globale est une conséquence de cette estimation.

□

Par symétrie, nous réécrivons la formulation variationnelle pour le flux convectif comme suit : pour $\varphi \in C^\infty_0(\mathbb{R})$,

$$\int_{\mathbb{R}} \varphi(x)\rho(x)\,\partial_x S_\varepsilon[\rho]\,\mathrm{d}x = -\frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{(\varphi(x)-\varphi(y))(x-y)}{|x-y|(|x-y|+\varepsilon)}\rho(x)\rho(y)\,\mathrm{d}x\,\mathrm{d}y. \tag{7.13}$$

Cette nouvelle formulation nous permet d'obtenir l'estimation suivante, uniforme en ε

$$\left|\int_{\mathbb{R}} \varphi\rho^\varepsilon\,\partial_x S_\varepsilon[\rho^\varepsilon]\,\mathrm{d}x\right| \leq \frac{M^2}{2\pi}\|\varphi'\|_\infty. \tag{7.14}$$

en utilisant que φ est lipschitzienne et la conservation de la masse (7.11). Nous introduisons la famille :

$$m^\varepsilon(t,x) := \int_{\mathbb{R}} K^\varepsilon(x-y)\rho^\varepsilon(x)\rho^\varepsilon(y)\,\mathrm{d}x\,\mathrm{d}y, \qquad \text{où} \qquad K^\varepsilon(x) = \frac{|x|}{|x|+\varepsilon}. \tag{7.15}$$

Dans la suite, nous considérons $\rho^\varepsilon(\cdot,t)$ et $m^\varepsilon(\cdot,t)$ comme des mesures dépendant du temps $\rho^\varepsilon(t)$ et $m^\varepsilon(t)$.

Lemme 7.2.7 *Les familles $\{\rho^\varepsilon(t)\}_{\varepsilon>0}$ et $\{m^\varepsilon(t)\}_{\varepsilon>0}$ sont tendues localement uniformément en temps et $\{\rho^\varepsilon(t)\}_{\varepsilon>0}$ est équicontinue en t pour la topologie de la convergence étroite.*

Preuve. Nous commençons par démontrer l'équicontinuité.

- **Équicontinuité de $\rho^\varepsilon(t)$ pour la topologie de la convergence étroite.** Soit $\varphi \in W^{2,\infty}(\mathbb{R})$. En utilisant (7.5), puis des intégrations par parties, nous obtenons :

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} \varphi\rho^\varepsilon\,\mathrm{d}x &= \int_{\mathbb{R}} \varphi\partial_t\rho^\varepsilon\,\mathrm{d}x \\
&\overset{(7.5)}{=} -\int_{\mathbb{R}} \varphi\ \partial_x(\chi\rho^\varepsilon\,\partial_x S_\varepsilon[\rho^\varepsilon] - \partial_x\rho^\varepsilon)\,\mathrm{d}x \\
&\overset{IPP}{=} \int_{\mathbb{R}} \partial_x\varphi(\chi\rho^\varepsilon\,\partial_x S_\varepsilon[\rho^\varepsilon] - \partial_x\rho^\varepsilon)\,\mathrm{d}x \\
&\overset{IPP}{=} \int_{\mathbb{R}} \rho^\varepsilon\left(\chi\,\partial_x\varphi\,\partial_x S_\varepsilon[\rho^\varepsilon] + \partial_x^2\varphi\right)\mathrm{d}x \\
&= \int_{\mathbb{R}} \chi\rho^\varepsilon\,\partial_x\varphi\,\partial_x S_\varepsilon[\rho^\varepsilon]\,\mathrm{d}x + \int_{\mathbb{R}} \partial_x^2\varphi\rho^\varepsilon\,\mathrm{d}x.
\end{aligned}$$

  Finalement, par (7.14), par conservation de la masse (7.11) et positivité de ρ^ε (7.10),

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}} \varphi\rho^\varepsilon\,\mathrm{d}x\right| \le \frac{\chi M^2}{2\pi}\|\varphi''\|_\infty + M\|\varphi''\|_\infty =: c\|\varphi''\|_\infty,$$

  où la constante c ne dépend ni de t, ni de ε. Cela nous donne l'équicontinuité dans $\left(W^{2,\infty}(\mathbb{R})\right)'$:

$$\left|\int_{\mathbb{R}} \varphi\rho^\varepsilon(t,x)\,\mathrm{d}x - \int_{\mathbb{R}} \varphi\rho^\varepsilon(x,s)\,\mathrm{d}x\right| \le C(\varphi)|t-s|. \tag{7.16}$$

  Afin d'obtenir l'équicontinuité pour la topologie de la convergence étroite nous devons obtenir le même type de résultat pour toute fonction $\varphi \in C_b(\mathbb{R})$.

- **Propriété de tension uniforme.** Soit la fonction test $\varphi_R(x) = 1 - \beta\left(\frac{x^2}{R^2}\right)$ avec β décroissante, $\beta(r) = 1$ pour $0 \le r \le 1/2$, et $\beta(r) = 0$ pour $r \ge 1$. Nous l'injectons dans l'inégalité précédente et nous obtenons :

$$\int_{\mathbb{R}\setminus[-R,R]} \rho^\varepsilon\,\mathrm{d}x \le \int_{\mathbb{R}\setminus\left[-\frac{R}{2},\frac{R}{2}\right]} \rho^\varepsilon\,\mathrm{d}x + \frac{Ct}{R^2}, \tag{7.17}$$

  ce qui donne immédiatement que la suite est tendue uniformément sur tout intervalle de temps borné. Puisque $|m^\varepsilon| \le M\rho^\varepsilon$ le résultat vaut également pour m^ε.

  Nous en déduisons (7.16) pour des fonctions $\varphi \in C_b(\mathbb{R})$ de la manière suivante : (7.17) montre que pour tout $\delta > 0$, pour tout $t < \infty$, il existe une constante $R_{t,\delta} < \infty$ telle que :

$$\left|\int_{\mathbb{R}\setminus[-R_{t,\delta},R_{t,\delta}]} \rho^\varepsilon(t,x)\varphi(x)\,\mathrm{d}x\right| \le \|\varphi\|_\infty.$$

Sur le segment $[-R_{t,\delta}, R_{t,\delta}]$ nous procédons par approximation : pour tout $\delta > 0$, il existe $\varphi_\delta \in W^{2,\infty}(\mathbb{R})$ telle que $\|\varphi - \varphi_\delta\|_{L^\infty([-R_{t,\delta},R_{t,\delta}])} \leq \delta$. Puisque ρ^ε est borné uniformément, nous obtenons :

$$\left| \int_{[-R_{t,\delta},R_{t,\delta}]} \varphi\rho^\varepsilon(t,x)\,\mathrm{d}x - \int_{[-R_{t,\delta},R_{t,\delta}]} \varphi\rho^\varepsilon(x,s)\,\mathrm{d}x \right| \leq 2\delta M + C(\varphi_\delta)|t-s|,$$

Finalement,

$$\left| \int_{\mathbb{R}} \varphi\rho^\varepsilon(t,x)\,\mathrm{d}x - \int_{\mathbb{R}} \varphi\rho^\varepsilon(x,s)\,\mathrm{d}x \right| \leq 2\delta M + C(\varphi_\delta)|t-s|,$$

Comme $\displaystyle\int_{\mathbb{R}} \rho^\varepsilon\,\mathrm{d}x = M$, ceci montre l'équicontinuité pour la convergence étroite de $\rho^\varepsilon(t)$.

□

Nous sommes alors en mesure d'appliquer le théorème de Prokhorov : il existe des mesures finies positives dépendant du temps $\rho(t)$ et $m(t)$ telles que, à extraction près, $\rho^\varepsilon(t)$ converge vers $\rho(t)$ étroitement et localement uniformément en t lorsque ε tend vers 0, *i.e.* :

$$\forall \varphi \in C_b(\mathbb{R}), \quad \lim_{\varepsilon\to 0} \int_{\mathbb{R}} \varphi(x)\rho^\varepsilon(t,x)\,\mathrm{d}x = \int_{\mathbb{R}} \varphi(x)\rho(t,x)\,\mathrm{d}x,$$

uniformément sur $[0,T] \subset \mathbb{R}$, $\forall T > 0$. De plus, pour toute fonction $\varphi \in C_b([t_1,t_2]\times\mathbb{R})$ nous avons la convergence suivante :

$$\lim_{\varepsilon\to 0} \int_{t_1}^{t_2} \int_{\mathbb{R}} \varphi(t,x) m^\varepsilon(t,x)\,\mathrm{d}x\,\mathrm{d}t = \int_{t_1}^{t_2} \int_{\mathbb{R}} \varphi(t,x) m(t,x)\,\mathrm{d}x\,\mathrm{d}t. \tag{7.18}$$

Finalement, pour tout $T > 0$, $\rho^\varepsilon(t)$ tend vers $\rho(t)$ étroitement, uniformément par rapport à $t \in [0,T]$. Par conséquent $\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)$ converge vers $\rho(t,x)\rho(t,y)$ étroitement, uniformément par rapport à $t \in [0,T]$ dans $\mathcal{M}_+^1(\mathbb{R}\times\mathbb{R})$. Nous en concluons que, pour toute fonction $\varphi \in C_b(\mathbb{R}\times\mathbb{R})$,

$$\lim_{\varepsilon\to 0} \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x,y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x,y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{7.19}$$

Cependant, étant donné que le noyau limite K défini par :

$$K(x) = \lim_{\varepsilon\to 0} K^\varepsilon(x) = \begin{cases} 1 & \text{si} \quad x \neq 0, \\ 0 & \text{si} \quad x = 0, \end{cases}$$

est discontinu, nous ne pouvons pas passer à la limite dans le membre de droite de (7.15), et nous devons introduire la mesure de défaut

$$\nu(t,x) = m(t,x) - \int_{\mathbb{R}} K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}y. \tag{7.20}$$

Le support atomique de la mesure $\rho(t)$ sera noté dans la suite

$$S_{at}(\rho(t)) := \{a \in \mathbb{R} \,:\, \rho(t)(\{a\}) > 0\}.$$

C'est un ensemble au plus dénombrable.

Lemme 7.2.8 (POUPAUD '02) *La mesure de défaut est positive et vérifie*

$$\nu(t,x) \leq \sum_{a\in S_{at}(\rho(t))} \Big(\rho(t)(\{a\})\Big)^2 \delta(x-a).$$

Preuve.

- Pour toute fonction test $\varphi \in C_b(\mathbb{R}\times\mathbb{R})$, nous avons

$$\big|\varphi(x,y)-\varphi(x,x)\big|\,\big|K^\varepsilon(x-y)-K(x-y)\big| \le C\min(\varepsilon,|x-y|),$$

où C est une constante dépendant de φ. Ainsi,

$$(\varphi(x,y)-\varphi(x,x))\,K^\varepsilon(x-y) \longrightarrow (\varphi(x,y)-\varphi(x,x))\,K(x-y),$$

uniformément, et par conséquent, d'après (7.19)

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\int_{\mathbb{R}}(\varphi(x,y)-\varphi(x,x))\,K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}}\int_{\mathbb{R}}(\varphi(x,y)-\varphi(x,x))\,K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y.$$

En utilisant les définitions de m (7.15) et ν (7.20) nous calculons :

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}}\varphi(x,y)K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\\
=\;&\int_{\mathbb{R}}\int_{\mathbb{R}}[\varphi(x,y)-\varphi(x,x)]\,K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\\
&\qquad+\int_{\mathbb{R}}\int_{\mathbb{R}}\varphi(x,x)K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y,\\
\overset{(7.15)}{=}\;&\int_{\mathbb{R}}\int_{\mathbb{R}}[\varphi(x,y)-\varphi(x,x)]\,K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y+\int_{\mathbb{R}}\varphi(x,x)m^\varepsilon(t,x)\,\mathrm{d}x,\\
\overset{(7.20)}{\longrightarrow}\;&\int_{\mathbb{R}}\int_{\mathbb{R}}[\varphi(x,y)-\varphi(x,x)]\,K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y\\
&\qquad+\int_{\mathbb{R}}\varphi(x,x)\left(\nu(t,x)+\int_{\mathbb{R}}K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}y\right)\mathrm{d}x.
\end{aligned}$$

Finalement, nous obtenons la convergence suivante :

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\int_{\mathbb{R}}\varphi(x,y)K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}}\int_{\mathbb{R}}\varphi(x,y)K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y+\int_{\mathbb{R}}\varphi(x,x)\nu(t,x)\,\mathrm{d}x. \quad (7.21)$$

Puisque $K^\varepsilon > 0$, le membre de droite est positif pour toute fonction test positive.
Pour montrer la positivité de la mesure de défaut ν, nous choisissons $\varphi(x,y)=\psi(x)\eta(R(x-y))\ge 0$ avec $\psi\in C(\mathbb{R})$ une fonction positive quelconque, et $\eta\in C_0(\mathbb{R})$ une fonction positive à support compact telle que $\eta(0)=1$ et R une constante. Nous injectons cette fonction test dans le premier terme du membre de droite de (7.21). Puisque $K(x-y)\eta(R(x-y))$ est borné et converge simplement vers zéro, ce terme tend vers zéro lorsque R tend vers l'infini par le théorème de convergence dominéede Lebesgue. Il en retourne que

$$\int_{\mathbb{R}}\psi(x)\nu(t,x)\,\mathrm{d}x\ge 0$$

pour toute fonction positive ψ, comme annoncé.

- Nous démontrons maintenant la seconde partie du lemme. Pour cela, notons tout d'abord que $K^\varepsilon\le 1$. Ainsi, pour toute fonction test positive $\varphi\in C_b(\mathbb{R}\times\mathbb{R})$,

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\varphi(x,y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\ge\int_{\mathbb{R}}\int_{\mathbb{R}}\varphi(x,y)K^\varepsilon(x-y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y.$$

En passant à la limite dans cette inégalité, nous obtenons :

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x,y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y \geq \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x,y)K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y + \int_{\mathbb{R}} \varphi(x,x)\nu(t,x)\,\mathrm{d}x.$$

Remarquons que $K(x-y) = 1-\chi_D(x,y)$ où χ_D désigne la fonction caractéristique de la diagonale de $\mathbb{R}\times\mathbb{R}$. De plus, puisque

$$\chi_D(x,y)\rho(t,x)\rho(t,y) = \sum_{a\in S_{at}(\rho(t))} \rho(t)(\{a\})^2\delta(x-a)\delta(y-a),$$

l'inégalité précédente donne

$$\sum_{a\in S_{at}(\rho(t))} \rho(t)(\{a\})^2\delta(x-a)\delta(y-a) \geq \int_{\mathbb{R}} \varphi(x,x)\nu(t,x)\,\mathrm{d}x$$

ce qui conduit au résultat désiré.

□

La limite de ρ^ε lorsque ε tend vers 0 est caractérisée par la paire (ρ,ν) dont les propriétés sont détaillées dans la définition suivante.

Définition 7.2.9 ([Pou02]) *Pour tout intervalle $I \subset \mathbb{R}$ borné, l'ensemble des mesures dépendant du temps à défauts diagonaux est défini par*

$$\mathcal{DM}^+(I;\mathbb{R}) = \left\{ \begin{array}{l} (\rho,\nu) \ : \ \rho(t) \in \mathcal{M}_+^1(\mathbb{R})\ \forall t\in I,\ \nu\in\mathcal{M}(I\times\mathbb{R}), \\ \rho \textit{ est continue par rapport à } t \textit{ pour la topologie de la convergence étroite} \\ \nu \textit{ est positive}, \\ \nu(t,x) \leq \displaystyle\sum_{a\in S_{at}(\rho(t))} (\rho(t)(\{a\}))^2\delta(x-a). \end{array} \right\}$$

où $\mathcal{M}$ désigne l'espace des mesures de Radon et $\mathcal{M}_+^1$ le sous–espace des mesures positives et bornées.

7.2.4 Solutions mesures généralisées

Nous voulons à présent passer à la limite dans le modèle régularisé de Patlak–Keller–Segel (7.5). Pour cela, nous injectonsl'identité suivante dans (7.13) :

$$\frac{(\varphi(x)-\varphi(y))(x-y)}{|x-y|(|x-y|+\varepsilon)} = K^\varepsilon(x-y)\,\partial_x\varphi(x) + L^\varepsilon(\varphi)(x,y), \tag{7.22}$$

où

$$L^\varepsilon(\varphi)(x,y) = \frac{\Big(\varphi(x)-\varphi(y)-(x-y)\ \partial_x\varphi(x)\Big)(x-y)}{|x-y|(|x-y+\varepsilon)},$$

qui converge uniformément vers la fonction continue $L^0(\varphi)(x,y)$,

$$L^0(\varphi)(x,y) = \frac{(\varphi(t,x)-\varphi(t,y))(x-y)}{|x-y|^2} - \partial_x\varphi(t,x)K(x-y), \tag{7.23}$$

pour toute fonction test $\varphi \in C_b^1(\mathbb{R})$.
Pour tout intervalle de temps $(0,T)$ et pour toute fonction test $\varphi \in C_b^1((0,T)\times\mathbb{R})$, nous calculons :

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \varphi(t,x)\rho^\varepsilon(t,x)\,\partial_x S_\varepsilon[\rho^\varepsilon](t,x)\,\mathrm{d}x\,\mathrm{d}t \\
\overset{(7.13)}{=}\quad & -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{(\varphi(t,x)-\varphi(t,y))(x-y)}{|x-y|(|x-y|+\varepsilon)}\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \\
\overset{(7.22)}{=}\quad & -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} K^\varepsilon(x-y)\,\partial_x\varphi(x)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \\
&\qquad -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} L^\varepsilon(\varphi)(x,y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \\
\overset{(7.15)}{=}\quad & -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(x)m^\varepsilon(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} L^\varepsilon(\varphi)(x,y)\rho^\varepsilon(t,x)\rho^\varepsilon(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t.
\end{aligned}$$

Ainsi, (7.18) montre que

$$\begin{aligned}
\lim_{\varepsilon\to 0}\int_0^T \int_{\mathbb{R}} \varphi(t,x)\rho^\varepsilon(t,x)\,\partial_x S_\varepsilon[\rho^\varepsilon](t,x)\,\mathrm{d}x\,\mathrm{d}t = \\
-\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(x)m(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} L^0(\varphi)(x,y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t.
\end{aligned}$$

En utilisant (7.20) et (7.23), cette limite se réécrit :

$$\begin{aligned}
& -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(x)m(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} L^0(\varphi)(x,y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \\
= \quad & -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\nu(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} \partial_x\varphi(t,x)K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t \\
& \qquad -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{(\varphi(t,x0-\varphi(t,y))(x-y)}{|x-y|^2}\rho(t,x)\rho(t,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t \\
& \qquad +\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} \partial_x\varphi(t,x)K(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t, \\
= \quad & -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\nu(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} (\varphi(t,x)-\varphi(t,y))\,\frac{x-y}{|x-y|^2}\rho(t,x)\rho(t,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

En résumé, nous obtenons le résultat de convergence suivant :

$$\begin{aligned}
&\lim_{\varepsilon\to 0} -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(x)m(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} L^0(\varphi)(x,y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \\
&\quad = -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\nu(t,x)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} (\varphi(t,x)-\varphi(t,y))\,\mathcal{K}(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t,
\end{aligned}$$

où

$$\mathcal{K}(x) = \begin{cases} \frac{1}{x} & \text{for} \quad x \neq 0, \\ 0 & \text{for} \quad x = 0. \end{cases}$$

Par conséquent, à extraction près, $\rho^\varepsilon\, \partial_x S_\varepsilon[\rho^\varepsilon]$ converge vers $j[\rho, v]$ au sens des distributions, où le flux limite est défini pout tout $\varphi \in C_b^1((0,T)\times\mathbb{R})$ par

$$\begin{aligned}\int_0^T \int_{\mathbb{R}} \varphi(t,x) j[\rho,v]\,\mathrm{d}x\,\mathrm{d}t \\ = -\frac{1}{2\pi}\int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}} (\varphi(t,x)-\varphi(t,y))\mathcal{K}(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t - \frac{1}{2\pi}\int_0^T\int_{\mathbb{R}} v(t,x)\,\partial_x\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned} \tag{7.24}$$

Cela démontre le théorème principal de ce paragraphe :

Théorème 7.2.3 *Pour tout $T > 0$, la solution ρ^ε de (7.5)–(7.7) converge à extraction près étroitement et uniformément en temps vers une mesure dépendant du temps $\rho(t)$. Par ailleurs, il existe une mesure de défaut $v(t)$ telle que $(\rho, v) \in \mathcal{DM}^+((0,T);\mathbb{R})$ soit une solution généralisée de*

$$\partial_t\rho + \partial_x(\chi j[\rho,v] - \partial_x\rho) = 0, \tag{7.25}$$

au sens où le flux convectif $j[\rho, v]$ est donné par (7.24) et que (7.25) soit vérifiée au sens des distributions. La condition initiale $\rho(t=0) = \rho_0$ est vérifiée au sens de la continuité étroite.

Remarque 15 *Remarquons que si $v = 0$,* i.e. *il n'y a pas de point de concentration, l'équation (7.25) se réduit au modèle classique.*

7.2.5 Système après blow–up

Cette section est consacrée à l'obtention d'une formulation forte pour (7.25). A cette fin, nous faisons l'hypothèse que la densité de cellules se décompose sous la forme

$$\rho = \overline{\rho} + \hat{\rho}, \tag{7.26}$$

où

$$\hat{\rho}(t,x) = \sum_{\ell\in L} M_\ell(t)\,\delta_\ell(t,x),$$

avec $\delta_\ell(t,x) = \delta(x - x_\ell(t))$ et $L \subset \mathbb{N}$, $\overline{\rho}$ est une fonction continue, C^1 par morceaux et les points de discontinuité de $\bar{\rho}$ coïncident avec les $(x_\ell(t))_{\ell\in L}$. De plus, nous supposons que, pour tout $\ell \in L$, $\overline{\rho}\,(x_\ell) = 0$. Nous supposons également que t varie dans un intervalle de temps où le support atomique de ρ consiste en des trajectoires régulières $x_\ell(t)$ portant des poids réguliers $M_\ell(t)$. Ceci constitue une façon parmi d'autres de décrire le système. À partir de cette description, la formulation forte obtenue est la suivante : .

$$\partial_t\,\overline{\rho} + \chi\,\partial_x(\overline{\rho}\;\partial_x S_0[\,\overline{\rho}\,]) - \sum_{\ell\in L}\frac{\chi}{\pi}\,\partial_x\left(\overline{\rho}\frac{M_\ell}{x-x_\ell}\right) - \partial_{xx}\overline{\rho} = 0, \tag{7.27a}$$

$$\dot{M}_\ell = \left(\frac{\chi}{\pi}M_\ell + 1\right)\Big(\;\partial_x\,\overline{\rho}\,(t,x_\ell^+) - \;\partial_x\,\overline{\rho}\,(t,x_\ell^-)\Big), \tag{7.27b}$$

$$\dot{x}_\ell = \chi\,\partial_x S_0[\,\overline{\rho}\,](x_\ell) - \frac{\chi}{\pi}\sum_{m\neq\ell}\frac{M_m}{x_\ell - x_m}. \tag{7.27c}$$

Pour la décomposition (7.26), puisque $(\rho, v) \in \mathcal{DM}^+((0,T);\mathbb{R})$,

$$v(t,x) = \sum_{\ell\in L} v_\ell(t)\delta_\ell(t,x), \tag{7.28}$$

avec v_ℓ positive vérifiant $v_\ell \le M_\ell^2$. Nous pouvons alors réécrire le flux convectif défini de manière faible en (7.24) de manière forte comme suit :

$$j[\rho,v] = \overline{\rho}\;\partial_x\Big(S_0[\overline{\rho}+\hat{\rho}]\Big) + \sum_\ell M_\ell\delta_\ell\,\partial_x\left(S_0\left[\overline{\rho} + \sum_{m\neq\ell} M_m\delta_m\right] + \frac{1}{2\pi}\sum_\ell v_\ell\,\partial_x\delta_\ell\right), \tag{7.29}$$

où S_0 est défini presque partout par

$$S_0[\rho](t,x) = -\frac{1}{\pi}\int_{\mathbb{R}} \log|x-y|\rho(t,\,\mathrm{d}y).$$

Une fois cela montrer, nous réécrivons (7.25) afin d'obtenir la fomulation (7.27).

Calculs pour la réécriture de $j[\rho,v]$. Nous reprenons la définition (7.24) de j :

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}} \varphi(t,x) j[\rho,v]\,\mathrm{d}x\,\mathrm{d}t\\
=\;& -\frac{1}{2\pi}\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}} (\varphi(t,x)-\varphi(t,y))\mathcal{K}(x-y)\rho(t,x)\rho(t,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t + \frac{1}{2\pi}\int_0^T\int_{\mathbb{R}} v(t,x)\,\partial_x\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t,\\
=\;& \int_0^T\int_{\mathbb{R}} \varphi(t,x)\rho(t,\,\mathrm{d}x)\partial_x\Big(S_0[\rho]\Big)\,\mathrm{d}t + \frac{1}{2\pi}\int_0^T\int_{\mathbb{R}} v(t,\,\mathrm{d}x)\,\partial_x\varphi(t,x)\,\mathrm{d}t.
\end{aligned}$$

Nous avons

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}} \varphi(t,x)\rho(t,\,\mathrm{d}x)\partial_x\Big(S_0[\rho]\Big)\,\mathrm{d}t\\
=\;& \int_0^T\int_{\mathbb{R}} \varphi(t,x)\Big(\bar\rho+\hat\rho\Big)(t,\,\mathrm{d}x)\partial_x\Big(S_0[\bar\rho+\hat\rho]\Big)\,\mathrm{d}t,\\
=\;& \int_0^T\int_{\mathbb{R}} \varphi(t,x)\bar\rho(t,\,\mathrm{d}x)\partial_x\Big(S_0[\bar\rho+\hat\rho]\Big)\,\mathrm{d}t + \int_0^T\int_{\mathbb{R}} \varphi(t,x)\hat\rho(t,\,\mathrm{d}x)\partial_x\Big(S_0[\bar\rho+\hat\rho]\Big)\,\mathrm{d}t.
\end{aligned}$$

L'hypothèse de structure (7.26)–(7.28) implique

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}} \varphi(t,x)\rho(t,\,\mathrm{d}x)\partial_x\Big(S_0[\rho]\Big)\,\mathrm{d}t\\
=\;& \frac{1}{\pi}\int_0^T\int_{\mathbb{R}} \varphi(t,x)\hat\rho(t,\,\mathrm{d}x)\int_{\mathbb{R}}\mathcal{K}(x-y)\,(\overline\rho+\hat\rho)\,(t,\,\mathrm{d}y)\,\mathrm{d}t\\
=\;& \frac{1}{\pi}\int_0^T\int_{\mathbb{R}} \varphi(t,x)\hat\rho(t,\,\mathrm{d}x)\int_{\mathbb{R}}\mathcal{K}(x-y)\left(\overline\rho+\sum_{m\in L}M_m\delta_m\right)(t,\,\mathrm{d}y)\,\mathrm{d}t\\
=\;& \frac{1}{\pi}\int_0^T\int_{\mathbb{R}} \varphi(t,x)\sum_{\ell\in L}M_\ell\delta_\ell(\,\mathrm{d}x)\int_{\mathbb{R}}\mathcal{K}(x-y)\left(\overline\rho+\sum_{m\neq\ell}M_\ell\delta_\ell\right)(t,\,\mathrm{d}y)\,\mathrm{d}t\\
=\;& \frac{1}{\pi}\int_0^T\int_{\mathbb{R}} \varphi(t,x)\sum_{\ell\in L}M_\ell\delta_\ell(x)\,\partial_x S_0\left[\overline\rho+\sum_{m\neq\ell}M_m\delta_m\right](t,\,\mathrm{d}x)\,\mathrm{d}t.
\end{aligned}$$

Ainsi

$$\begin{aligned}
\int_0^T\int_{\mathbb{R}} \varphi(t,x)j[\rho,v]\,\mathrm{d}x\,\mathrm{d}t =\;& \int_0^T\int_{\mathbb{R}} \varphi(t,x)\,\overline\rho(t,x)\;\partial_x\Big(S_0[\overline\rho+\hat\rho]\Big)(t,\,\mathrm{d}x)\,\mathrm{d}t\\
&+\frac{1}{\pi}\int_0^T\int_{\mathbb{R}} \varphi(t,x)\sum_{\ell\in L}M_\ell\delta_\ell(x)\,\partial_x S_0\left[\overline\rho+\sum_{m\neq\ell}M_m\delta_m\right](t,\,\mathrm{d}x)\,\mathrm{d}t + \frac{1}{2\pi}\int_0^T\sum_{\ell\in L}v_\ell\,\partial_x\varphi(t,x_\ell)\,\mathrm{d}t.
\end{aligned}$$

pour toute fonction test φ, d'où la réécriture du flux (7.29).

Réécriture de l'équation (7.25)**.** Nous pouvons réécrire (7.25) à l'aide de la reformulation (7.29) de $j[\rho, v]$ sous l'hypothèse de structure (7.26)–(7.28). Pour toute fonction test $\varphi \in \mathcal{D}((0,T)\times\mathbb{R})'$:

$$\underbrace{\int_0^T \int_{\mathbb{R}} \varphi(t,x)\partial_t \rho(t,\,\mathrm{d}x)\,\mathrm{d}t}_{=:\,\boxed{1}} + \underbrace{\int_0^T \int_{\mathbb{R}} \varphi(t,x)\;\partial_x\left(\chi j[\rho,v] - \partial_x\rho\right)(t,\,\mathrm{d}x)\,\mathrm{d}t}_{=:\,\boxed{2}} = 0$$

Nous traitons les différents termes séparément :

$$\begin{aligned}
\boxed{1} &= \int_0^T \int_{\mathbb{R}} \varphi(t,x)\partial_t\,[\,\overline{\rho}+\hat{\rho}\,](t,\,\mathrm{d}x)\,\mathrm{d}t \\
&= \int_0^T \int_{\mathbb{R}} \varphi\left(\partial_t\,\overline{\rho} + \sum_{\ell\in L}\left(\dot{M}_\ell\,\delta_\ell - M_\ell\,\dot{x}_\ell\;\partial_x\delta_\ell\right)\right)(t,\,\mathrm{d}x)\,\mathrm{d}t, \\
\boxed{2} &= -\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\,(j[\rho,v] - \partial_x\rho)\,(t,\,\mathrm{d}x)\,\mathrm{d}t \\
&= -\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\left(\chi\overline{\rho}\;\partial_x\Big(S_0[\overline{\rho}+\hat{\rho}]\Big) + \chi\sum_{\ell\in L} M_\ell\delta_\ell\;\partial_x\left(S_0\left[\,\overline{\rho}+\sum_{m\neq\ell}M_m\delta_m\right] + \frac{\chi}{2\pi}\sum_{\ell\in L}\nu_\ell\;\partial_x\delta_\ell\right)\right)(t,\,\mathrm{d}x)\,\mathrm{d}t \\
&\qquad\qquad + \int_0^T \int_{\mathbb{R}} \partial_x\varphi\,\partial_x\rho\,\mathrm{d}x\,\mathrm{d}t \\
&\overset{IPP}{=} \underbrace{-\chi\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\,(\,\overline{\rho}\;\partial_x S_0[\,\overline{\rho}+\hat{\rho}\,])\,(t,\,\mathrm{d}x)\,\mathrm{d}t}_{=:\,\boxed{2.1}} \quad \underbrace{-\chi\int_0^T \sum_{\ell\in L} \partial_x\varphi(x_\ell)M_\ell\;\partial_x S_0\left[\,\overline{\rho}+\sum_{m\neq\ell}M_m\delta_m\right](x_\ell)\,\mathrm{d}t}_{=:\,\boxed{2.2}} \\
&\qquad \underbrace{-\int_0^T \int_{\mathbb{R}} \frac{\chi}{2\pi}\,\partial_x\varphi(t,x)\sum_{\ell\in L}\nu_\ell\,\partial_x\delta_\ell(t,\,\mathrm{d}x)\,\mathrm{d}t}_{=:\,\boxed{2.3}} + \underbrace{\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\,\partial_x\,(\,\overline{\rho}+\hat{\rho}\,)\,(t,\,\mathrm{d}x)\,\mathrm{d}t}_{=:\,\boxed{2.4}}
\end{aligned}$$

À nouveau, nous calculons chaque terme séparemment. Notons que pour le calcul de $\boxed{2.1}$, l'hypothèse $\bar{\rho}(x_\ell) = 0$ pour tout $\ell \in L$ est essentielle. Nous renvoyons à la remarque 17 pour plus de détails.

$$\begin{aligned}
\boxed{2.1} &= -\chi\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\,(\,\overline{\rho}\;\partial_x S_0[\,\overline{\rho}+\hat{\rho}\,])\,(t,\,\mathrm{d}x)\,\mathrm{d}t \\
&= -\chi\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\overline{\rho}\left(\partial_x S_0[\,\overline{\rho}\,] - \frac{1}{\pi}\sum_{\ell\in L}\frac{M_\ell}{x-x_\ell}\right)(t,\,\mathrm{d}x)\,\mathrm{d}t
\end{aligned}$$

$$= -\chi \int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\overline{\rho}(t,x)\ \partial_x S_0[\,\overline{\rho}\,](t,\mathrm{d}x)\,\mathrm{d}t + \frac{\chi}{\pi}\int_0^T \int_{\mathbb{R}} \partial_x\varphi(t,x)\overline{\rho}(t,x)\sum_{\ell\in L}\frac{M_\ell}{x-x_\ell}\,\mathrm{d}x\,\mathrm{d}t$$

$$\overset{IPP}{=} \chi\int_0^T\int_{\mathbb{R}}\varphi(t,x)\,\partial_x\left(\overline{\rho}\ \partial_x S_0[\,\overline{\rho}\,]\right)(t,\mathrm{d}x)\,\mathrm{d}t + \frac{\chi}{\pi}\int_0^T\sum_{\ell\in L}\left[-\int_{\mathbb{R}}\varphi\,\partial_x\left(\overline{\rho}\,\frac{M_\ell}{x-x_\ell}\right) - \lim_{\substack{x\to x_\ell\\ x>x_\ell}}\varphi M_\ell\frac{\overline{\rho}}{x-x_\ell} + \lim_{\substack{x\to x_\ell\\ x<x_\ell}}\varphi M_\ell\frac{\overline{\rho}}{x-x_\ell}\right]\mathrm{d}t$$

$$= \chi\int_0^T\int_{\mathbb{R}}\varphi(t,x)\,\partial_x\left(\overline{\rho}\ \partial_x S_0[\,\overline{\rho}\,]\right)(t,\mathrm{d}x)\,\mathrm{d}t + \frac{\chi}{\pi}\int_0^T\sum_{\ell\in L}\left[-\int_{\mathbb{R}}\varphi\,\partial_x\left(\overline{\rho}\,\frac{M_\ell}{x-x_\ell}\right) - \varphi(t,x_\ell)M_\ell\left(\partial_x\overline{\rho}(t,x_\ell^+) - \partial_x\overline{\rho}(t,x_\ell^-)\right)\right]\mathrm{d}t$$

$$= \int_0^T\int_{\mathbb{R}}\varphi\left(\chi\,\partial_x(\overline{\rho}\ \partial_x S_0[\,\overline{\rho}\,]) - \sum_{\ell\in L}\frac{\chi}{\pi}\,\partial_x\left(\overline{\rho}\frac{M_\ell}{x-x_\ell}\right) - \sum_{\ell\in L}\frac{\chi}{\pi}\delta_\ell M_\ell\left(\partial_x\overline{\rho}(\cdot,x_\ell^+) - \partial_x\overline{\rho}\,(\cdot,x_\ell^-)\right)\right)(t,\mathrm{d}x)\,\mathrm{d}t$$

$$\boxed{2.2} = -\chi\int_0^T\sum_{\ell\in L}\partial_x\varphi(x_\ell)M_\ell\ \partial_x S_0\left[\overline{\rho} + \sum_{m\neq\ell}M_m\delta_m\right](x_\ell)\,\mathrm{d}t,$$

$$= -\chi\int_0^T\sum_{\ell\in L}\partial_x\varphi(x_\ell)M_\ell\left(-\frac{1}{\pi}\int_{\mathbb{R}}\frac{1}{x_\ell-y}\overline{\rho}(y)\,\mathrm{d}y + \frac{1}{\pi}\sum_{m\neq\ell}\frac{M_m}{x_\ell-x_m}\right)\mathrm{d}t,$$

$$= +\frac{\chi}{\pi}\int_0^T\sum_{\ell\in L}\partial_x\varphi(x_\ell)M_\ell\int_{\mathbb{R}}\frac{1}{x_\ell-y}\overline{\rho}(y)\,\mathrm{d}y\,\mathrm{d}t + \frac{\chi}{\pi}\int_0^T\sum_{\ell\in L}\partial_x\varphi(x_\ell)M_\ell\sum_{m\neq\ell}\frac{M_m}{x_\ell-x_m}\,\mathrm{d}t,$$

$$= \int_0^T\int_{\mathbb{R}}\varphi\left(\sum_{\ell\in L}\chi\,\partial_x\delta_\ell M_\ell\,\partial_x S_0[\,\overline{\rho}\,](x_\ell) - \sum_{\ell\in L}\frac{\chi}{\pi}\,\partial_x\delta_\ell M_\ell\sum_{m\neq\ell}\frac{M_m}{x_\ell-x_m}\right)\mathrm{d}x\,\mathrm{d}t.$$

$$\boxed{2.3} = -\int_0^T\int_{\mathbb{R}}\frac{\chi}{2\pi}\,\partial_x\varphi(t,x)\sum_{\ell\in L}\nu_\ell\,\partial_x\delta_\ell(t,\mathrm{d}x)\,\mathrm{d}t,$$

$$\overset{IPP}{=} \int_0^T\frac{\chi}{2\pi}\sum_{\ell\in L}\nu_\ell\Delta\varphi(x_\ell)\,\mathrm{d}t.$$

$$\boxed{2.4} = \int_0^T\int_{\mathbb{R}}\partial_x\varphi(t,x)\,\partial_x\left(\overline{\rho}+\hat{\rho}\,\right)(t,\mathrm{d}x)\,\mathrm{d}t,$$

$$= -\int_0^T\int_{\mathbb{R}}\varphi\Delta\overline{\rho}\,\mathrm{d}x\,\mathrm{d}t - \int_0^T\sum_{\ell\in L}\varphi(t,x_\ell)\left(\partial_x\overline{\rho}(t,x_\ell^+) - \partial_x\overline{\rho}(t,x_\ell^-)\right)\mathrm{d}t - \int_0^T\sum_{\ell\in L}\Delta\varphi(x_\ell)M_\ell\,\mathrm{d}t,$$

$$= \int_0^T\int_{\mathbb{R}}\varphi\left(-\Delta\,\overline{\rho} - \sum_{\ell\in L}\delta_\ell\left(\partial_x\,\overline{\rho}\,(\cdot,x_\ell^+) - \partial_x\,\overline{\rho}\,(\cdot,x_\ell^-)\right) - \sum_{\ell\in L}\Delta\delta_\ell M_\ell\right)(t,\mathrm{d}x)\,\mathrm{d}t.$$

En ordonnant les termes, nous obtenons

$$\begin{aligned}\partial_t \overline{\rho} \quad &+\chi\, \partial_x(\overline{\rho}\, \partial_x S_0[\,\overline{\rho}\,]) - \sum_{\ell\in L} \frac{\chi}{\pi}\, \partial_x\Big(\overline{\rho}\frac{M_\ell}{x-x_\ell}\Big) - \Delta\overline{\rho} \\ &+\sum_{\ell\in L} \delta_\ell \left(\dot{M}_\ell - \Big(\frac{\chi}{\pi}M_\ell + 1\Big)\big(\,\partial_x\,\overline{\rho}\,(t,x_\ell^+) - \,\partial_x\,\overline{\rho}\,(t,x_\ell^-)\big)\right) \\ &+\sum_{\ell\in L} \partial_x\delta_\ell M_\ell \left(-\dot{x}_\ell + \chi\,\partial_x S_0[\,\overline{\rho}\,](x_\ell) - \frac{\chi}{\pi}\sum_{m\neq\ell}\frac{M_m}{x_\ell - x_m}\right) \\ &+\sum_{\ell\in L} \Delta\delta_\ell\Big(\frac{\chi}{2\pi}\nu_\ell - M_\ell\Big) = 0.\end{aligned}$$

On identifie chaque ligne à 0 suivant les dérivées des masses de Dirac qui sont des distributions d'ordres différents. La dernière donne $\chi\nu_\ell = 2\pi M_\ell$ ce qui combiné à $\nu_\ell \leq M_\ell^2$ donne la condition $\dfrac{2\pi}{\chi} \leq M_\ell$. Les points d'accumulation doivent avoir une masse d'au moins 2π. Les autres lignes donnent le système (7.27), comme annoncé.

Remarque 16 *Très bizarrement, la diffusion "alimente" la masse de Dirac grâce au second terme du membre de droite de* (7.27)*. Nous ne sommes pas en mesure de l'expliquer pour le moment.*

Le lemme suivant n'est pour l'instant pas encore démontré. Cependant, beaucoup d'indications nous invitent à penser qu'il est correct. C'est pourquoi nous nous permettons de l'énoncer ici.

Conjecture 1 *Les solutions $\overline{\rho}$ de* (7.27) *s'annulent aux points d'accumulation $(x_\ell)_{\ell\in L}$.*

Remarque 17 *L'idée essentielle en faveur de ce lemme est qu'une telle condition est nécessaire à la conservation de la masse. De plus, nous verrons à la Remarque 25, que numériquement la partie régulière de la densité s'annule aux points d'accumulation.*

C'est le système (7.27), que nous allons simuler numériquement dans la suite de ce chapitre. La méthode que nous proposons étend le schéma détaillé dans [BCC08] obtenu pour le système de Patlak–Keller–Segel avant explosion à notre problème. Ce schéma est fondé sur une formulation en terme de flot de gradient de l'équation, proposée originellement par Otto [Ott01, JKO98], faisant intervenir pour cela une nouvelle structure Riemannienne associée non plus à la métrique usuelle L^2 mais à la métrique de Wasserstein.

Dans la section qui suit, nous reviendrons sur le schéma de JKO avant de l'adapter au problème de Patlak–Keller–Segel.

7.3 Le schéma de Jordan–Kinderlehrer–Otto

Dans [Ott01], Otto introduit une structure Riemannienne formelle munissant l'équation des milieux poreux d'une structure de "flot de gradient" par rapport à la métrique de Wasserstein pour les mesures. Dans l'interprétation d'Otto les géodésiques correspondent alors à la trace du transport optimal

$$\rho^t = ((1-t)\mathrm{Id} + t\nabla\varphi)\,\#\rho^0,$$

où $T = \nabla\varphi$ est la solution du problème de transport optimal de ρ^0 vers ρ^1 et où le symbole # désigne le 'push–forward' défini plus loin, voir Définition 7.3.2.

Dans [JKO98], les auteurs adaptent ces idées du point de vue numérique en introduisant un algorithme de plus grande descente ('steepest descent scheme') basé sur le transport optimal de mesures pour l'équation linéaire de Fokker–Planck. L'idée correspond à une version discrète du flot de gradient pour la métrique de Wasserstein.

Pour le lecteur peu familier avec l'interprétation de flot de gradient et le transport optimal, nous proposons d'en décrire la philosophie à la Section 7.3.1 puis de rappeler les définitions de la distance de Wasserstein et du transport optimal à la Section 7.3.2. En Section 7.3.3, nous adaptons le schéma au problème de Patlak–Keller–Segel. Enfin, nous verrons qu'en dimension $d = 1$, le schéma de Jordan–Kinderlehrer–Otto se réduit en fait à un changement de fonction inconnue.

7.3.1 Philosophie

Dans [Ott01], Otto munit l'équation des milieux poreux d'une nouvelle structure de flot de gradient. Il s'avère que cette nouvelle structure est pertinente physiquement : en effet, cette structure de flot de gradient permet de séparer les quantités énergétiques des quantités cinétiques. Nous nous contenterons dans ce qui suit de décrire la structure abstraite de flot de gradient et de donner l'argument physique en faveur de la nouvelle structure. Nous renvoyons à [Ott01] pour une description plus détaillée et rigoureuse.

Dans ce paragraphe, nous nous intéressons uniquement à l'équation de la chaleur :

$$\frac{\partial \rho}{\partial t} - \Delta \rho = 0. \tag{7.30}$$

Dans la suite, $\rho \geq 0$ sera considéré comme une fonction densité dépendant du temps, dans tout l'espace $\mathbb{R}^d$.

Structure abstraite de flot de gradient

Avant d'interpréter l'équation de la chaleur comme un flot de gradient, détaillons quelque peu cette notion. Soient

- ⋆ $\mathcal{M}$ une variété différentiable ;
- ⋆ g un tenseur métrique faisant de $(\mathcal{M}, g)$ une variété Riemannienne ;
- ⋆ et une fonctionnelle E définie sur $\mathcal{M}$.

On appelle le système dynamique sur $\mathcal{M}$ défini par l'équation autonome suivante

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\mathrm{grad}\ E|_\rho\,, \tag{7.31}$$

le *flot de gradient de E sur* $(\mathcal{M}, g)$. Le rôle du tenseur métrique g est essentiellement de 'transformer' la différentielle $\mathrm{diff}E$ de E, qui est un champ de vecteurs cotangents, en gradient $\mathrm{grad}E$ de E, qui est cette fois un champ de vecteurs tangents :

$$g(\mathrm{grad}E, s) = \mathrm{diff}E.s \text{ pour tout champ de vecteurs } s \text{ sur } \mathcal{M}.$$

Ainsi (7.31) peut se réécrire de la façon suivante

$$g_\rho\left(\frac{\mathrm{d}\rho}{\mathrm{d}t}, s\right) + \mathrm{diff}\ E|_\rho\,.s = 0 \text{ pour tout champ de vecteurs } s \text{ le long de } \rho. \tag{7.32}$$

L'indice ρ indique ici que le tenseur g peut être fonction du point courant ρ. La caractéristique principale de la formulation en flot de gradient est la décroissance de l'énergie le long des trajectoires :

$$\frac{\mathrm{d}}{\mathrm{d}t}E(\rho) = \mathrm{diff}\ E|_\rho\,.\frac{d\rho}{dt} = -g_\rho\left(\frac{d\rho}{dt}, \frac{d\rho}{dt}\right) \leq 0. \tag{7.33}$$

Deux interprétations de l'équation de la chaleur en termes de flot de gradient

Dans cette section nous rappelons la formulation 'traditionnelle' de l'équation de la chaleur en terme de flot de gradient, à savoir le flot de gradient L^2 . Parallèlement à cela, nous introduisons également l'interprétation due à Otto [Ott01].

L'équation (7.30) préserve à la fois la positivité de la densité ρ et la masse totale $\int \rho$. Pour les deux approches, la variété est la suivante :

$$\mathcal{M} = \left\{ \rho \,:\, \mathbb{R}^d \to \mathbb{R}^+, \int \rho = 1 \right\}.$$

Sans entrer dans les détails de la structure de la variété, l'espace tangent est défini comme suit :

$$T_\rho \mathcal{M} = \left\{ s \,:\, \mathbb{R}^d \to \mathbb{R}, \int s = 0 \right\}.$$

Les différences entre les deux approches apparaissent dans la définition du tenseur métrique. Chacune est basée sur une identification différente de l'espace tangent

$$T_\rho \mathcal{M} \cong \left\{ s \,:\, \mathbb{R}^d \to \mathbb{R}, \int s = 0 \right\} / \mathbb{R}.$$

L'identification est réalisée via une équation elliptique

$$-\Delta p = s \text{ pour l'approche traditionnelle} \tag{7.34}$$

ou

$$-\nabla \cdot (\rho \nabla p) = s \text{ pour l'approche d'Otto.} \tag{7.35}$$

Le quotient signifie qu'on identifie deux fonctions p ne différant que d'une constante additive. Les tenseurs métriques sont alors définis par :

$$g_\rho(s_1, s_2) = \int \nabla p_1 \cdot \nabla p_2 \text{ pour l'approche traditionnelle} \tag{7.36}$$

ou

$$g_\rho(s_1, s_2) = \int \rho \nabla p_1 \cdot \nabla p_2 \text{ pour l'approche d'Otto} \tag{7.37}$$

où p_i est lié à s_i via l'identification (7.34) (resp. (7.35)). Dans les deux cas on obtient :

$$g_\rho(s_1, s_2) = \int s_1 p_2. \tag{7.38}$$

Finalement, la fonctionnelle E est donnée par

$$E(\rho) = \frac{1}{2} \int \rho^2 \text{ pour l'approche traditionnelle} \tag{7.39}$$

ou

$$E(\rho) = \int \rho \ln \rho \text{ pour l'approche d'Otto,} \tag{7.40}$$

tandis que sa différentielle s'écrit

$$\mathrm{diff}E(\rho).s = \int \rho s \text{ dans l'approche traditionnelle} \tag{7.41}$$

ou

$$\mathrm{diff}E(\rho).s = \int (\ln \rho + 1)\, s \text{ dans l'approche d'Otto.} \tag{7.42}$$

Montrons à présent que l'équation de la chaleur correspond aux flots de gradients évoqués, pour les deux approches.

Approche traditionnelle. En utilisant (7.38) et (7.41), l'égalité (7.32) se réécrit

$$\int \frac{\partial \rho}{\partial t} p + \int \rho s = 0,$$

où p est lié à s via $-\Delta p = s$. Après substitution et intégration par parties, nous obtenons

$$\int \left(\frac{\partial \rho}{\partial t} - \Delta \rho \right) p = 0.$$

p pouvant être arbitrairement choisi, nous retrouvons l'équation de la chaleur.

Approche d'Otto. En utilisant (7.38) et (7.42), l'égalité (7.32) se réécrit

$$\int \frac{\partial \rho}{\partial t} p + \int (\ln \rho + 1)\, s = 0,$$

où p est lié à s via $-\nabla \cdot (\rho \nabla p) = s$. Après substitution nous obtenons

$$\int \frac{\partial \rho}{\partial t} p - \int (\ln \rho + 1)\, \nabla \cdot (\rho \nabla p) = 0.$$

Enfin, par intégration par parties, nous obtenons

$$\int \left(\frac{\partial \rho}{\partial t} - \Delta \rho \right) p = 0,$$

et nous retrouvons à nouveau l'équation de la chaleur.

Dans le cadre de l'approche traditionnelle, la métrique est plate ; le tenseur métrique g ne dépend pas de ρ et par conséquent, g est le produit scalaire sur l'espace des fonctions de moyenne nulle. Dans ce cas, la structure Riemannienne $(\mathcal{M}, g)$ possède la structure d'un sous–espace convexe d'un espace Euclidien. Pour l'autre approche, la structure est plus compliquée. Nous donnons au paragraphe suivant un argument physique en faveur de cette nouvelle structure. Pour un argumentaire plus mathématique, nous renvoyons au travail d'Otto [Ott01].

Un argument physique

Avant d'interpréter physiquement la notion de flot de gradient décrite ci–dessus, donnons quelques informations concernant la physique de l'équation de la chaleur (7.30). La fonction ρ décrit la densité de masse d'un gaz, par exemple, dans un milieu. La première hypothèse est celle de conservation de masse, qui se traduit par l'expression suivante

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0, \tag{7.43}$$

où le vecteur u décrit la vitesse moyenne du gaz. La seconde hypothèse est celle de la loi de Darcy

$$u = -M.\nabla p,$$

où la fonction p définie sur $\mathbb{R}^d$ décrit la pression du gaz et la matrice M décrit la mobilité du gaz dans le milieu. Typiquement M dépend de la perméabilité du milieu et de la viscosité du gaz. Supposons la perméabilité isotropique et homogène, de telle façon qu'après adimensionnement nous ayons $M = \mathrm{id}$, ce qui conduit à

$$u = -\nabla p. \tag{7.44}$$

La dernière hypothèse vient de la thermodynamique :

$$p = \frac{\delta E}{\delta \rho}, \tag{7.45}$$

où E désigne l'énergie libre et $\dfrac{\delta E}{\delta \rho}$ sa fonction dérivée par rapport à la variable ρ. Si l'énergie libre s'écrit sous la forme

$$E = \int e(\rho),$$

où la fonction $z \mapsto e(z)$ décrit comment la densité d'énergie libre e dépend de la densité ρ, l'équation (7.45) se réécrit

$$p = e'(\rho). \tag{7.46}$$

En combinant ces trois hypothèses (7.43), (7.44) et (7.46) nous obtenons

$$\frac{\partial \rho}{\partial t} - \Delta \pi(\rho) = 0, \tag{7.47}$$

où la fonction $z \mapsto \pi(z)$ décrit le comportement de la pression osmotique π par rapport à la densité ρ. Cette fonction est liée à la densité d'énergie libre via la formule

$$\pi(z) = ze'(z) - e(z). \tag{7.48}$$

En utilisant (7.48), l'équation (7.47) correspond à l'équation de la chaleur (7.30) si

$$\pi(z) = z,$$

ce qui est vrai si et seulement si

$$e(z) = z \ln z.$$

Dans ce sens, uniquement la structure introduite par Otto munit la fonctionnelle E d'une signification physique.

Le tenseur métrique possède dans ce cadre également une signification physique. Remarquons que la définition (7.37) de g peut être reformulée ainsi

$$g_\rho(s,s) = \inf \left\{ \int \rho |u|^2,\ u : \mathbb{R}^d \to \mathbb{R}^d,\ s + \nabla \cdot (\rho u) = 0. \right\}. \tag{7.49}$$

En effet, le minimiseur de la fonctionnelle quadratique du problème (7.49) vérifie $\displaystyle\int v \cdot u = 0$ pour tous champs de vecteurs v sur $\mathbb{R}^d$ avec $\nabla \cdot v = 0$, de telle sorte qu'il existe une fonction p de $\mathbb{R}^d$ telle que

$$u = -\nabla p.$$

Observons maintenant que la quantité $\displaystyle\int \rho |u|^2$ dans (7.49) a une interprétation physique : elle correspond au taux de dissipation de l'énergie cinétique par friction lorsque le gaz se meut à la vitesse u à travers le milieu. Ainsi $g_\rho(s,s)$ mesure le taux minimal de dissipation de l'énergie cinétique par friction nécessaire pour modifier la densité ρ d'un taux s. Cela nous amène à une jolie interprétation de l'équation (7.33), à savoir

$$\frac{\mathrm{d}}{\mathrm{d}t} E(\rho) = -g_\rho \left(\frac{d\rho}{dt}, \frac{d\rho}{dt} \right).$$

Définition 7.3.1 *La distance induite par la structure Riemannienne $d(\rho_0, \rho_1)^2$ est définie comme l'infimum de l'énergie $\displaystyle\int_0^1 \left| \frac{d\tilde{\rho}}{d\sigma} \right|^2 \mathrm{d}\sigma$ sur toutes les courbes*

$$\begin{array}{rcl} [0,1] & \to & \mathcal{M} \\ \sigma & \mapsto & \tilde{\rho}(\sigma), \end{array}$$

reliant ρ_0 à ρ_1 :

$$d(\rho_0, \rho_1)^2 = \inf \left\{ \int_0^1 \left| \frac{d\tilde{\rho}}{d\sigma} \right|^2 \mathrm{d}\sigma \ \middle|\ \begin{array}{rcl} [0,1] & \to & \mathcal{M} \\ \sigma & \mapsto & \tilde{\rho}(\sigma), \end{array},\ \tilde{\rho}(0) = \rho_0,\ \tilde{\rho}(1) = \rho_1 \right\}.$$

$\tilde{\rho}$ qui réalise l'infimum est la géodésique *reliant ρ_0 à ρ_1.*

Remarque 18 *On a aussi :*

$$d(\rho_0,\rho_1)^2 = \int_0^1 g_{\tilde\rho}\left(\frac{d\tilde\rho}{dt},\frac{d\tilde\rho}{dt}\right)\mathrm{d}\sigma,$$

où $\tilde\rho$ est la géodésique reliant ρ_0 à ρ_1.

Dans [Ott01], l'auteur montre que pour la structure décrite ci–avant, caractérisée par l'identification avec le plan tangent (7.35), cette distance correspond à *la distance de Wasserstein*. Nous reviendrons sur cette notion en Section 7.3.2.

Discrétisation naturelle pour le problème de flot de gradient (7.31)

Dans [JKO98], les auteurs donnent une discrétisation naturelle en temps pour le système dynamique (7.31). Ce schéma est celui qui est décrit originellement dans [Ott01] pour l'équation de Fokker–Plank et qui est connu sous le nom de schéma de Jordan–Kinderlehrer–Otto.

L'équation de flot de gradient (7.31)

$$\frac{d\rho}{dt} = -\mathrm{grad}\ E|_\rho\,,$$

admet la discrétisation en temps suivante. Soit $\Delta t > 0$, un pas de temps donné, nous considérons le schéma itératif suivant

$$\left.\begin{array}{l}\rho^{(k)} \text{ minimise}\\ \dfrac{1}{2\Delta t}d\left(\rho^{(k-1)},\rho\right)^2 + E(\rho)\\ \text{sur tous les } \rho\in\mathcal{M}\end{array}\right\}, \tag{7.50}$$

où d correspond à la distance induite sur $(\mathcal{M},g)$.

Redonnons ici l'argument de [Ott01] montrant que (7.50) est bien une discrétisation de l'équation de flot de gradient. À cette fin, nous dérivons la première variation du problème de minimisation (7.50). Soit

$$\begin{array}{rcl}[0,1] & \to & \mathcal{M}\\ \sigma & \mapsto & \tilde\rho^{(k)}\end{array}$$

la courbe d'énergie minimale reliant $\rho^{(k-1)}$ et $\rho^{(k)}$. Soit une variation $\rho_\varepsilon^{(k)}$ de $\rho^{(k)}$, à savoir une courbe $\varepsilon\mapsto\rho_\varepsilon^{(k)}\in\mathcal{M}$ avec $\rho_{\varepsilon=0}^{(k)}=\rho^{(k)}$. Soit finalement

$$\begin{array}{rcl}[0,1] & \to & \mathcal{M}\\ \sigma & \mapsto & \tilde\rho_\varepsilon^{(k)}\end{array}$$

une courbe reliant $\rho^{(k-1)}$ à $\rho_\varepsilon^{(k)}$ coïncidant avec $\tilde\rho^{(k)}$ pour $\varepsilon=0$. Nous résumons les notations sur le schéma ci–contre, Figure 7.2.

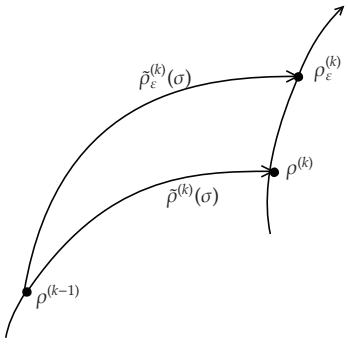


FIG. 7.2: Schéma : variation autour de la géodésique

Par définition du schéma et de la distance nous avons

$$\begin{array}{rcl}\dfrac{1}{2\Delta t}\displaystyle\int_0^1\left|\frac{d\tilde\rho^{(k)}}{d\sigma}\right|^2\mathrm{d}\sigma + E(\rho^{(k)}) & = & \dfrac{1}{2\Delta t}d\left(\rho^{(k-1)},\rho^{(k)}\right)^2 + E(\rho^{(k)})\\ & \le & \dfrac{1}{2\Delta t}d\left(\rho^{(k-1)},\rho_\varepsilon^{(k)}\right)^2 + E(\rho_\varepsilon^{(k)}),\\ & \le & \dfrac{1}{2\Delta t}\displaystyle\int_0^1\left|\frac{d\tilde\rho_\varepsilon^{(k)}}{d\sigma}\right|^2\mathrm{d}\sigma + E(\rho_\varepsilon^{(k)}).\end{array}$$

Il s'ensuit,

$$\begin{aligned}
0 &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\left(\frac{1}{2\Delta t}\int_0^1\left|\frac{d\tilde\rho_\varepsilon^{(k)}}{d\sigma}\right|^2 \mathrm{d}\sigma + E(\rho_\varepsilon^{(k)})\right),\\
&= \frac{1}{\Delta t}\int_0^1\left\langle\frac{d\tilde\rho^{(k)}}{d\sigma}, \frac{D}{d\sigma}\left.\frac{d\tilde\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle \mathrm{d}\sigma + \left\langle \mathrm{grad}\, E|_{\rho^{(k)}}, \left.\frac{d\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle,\\
&= \frac{1}{\Delta t}\int_0^1\left\{\frac{d}{d\sigma}\left\langle\frac{d\tilde\rho^{(k)}}{d\sigma}, \left.\frac{d\tilde\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle - \left\langle\underbrace{\frac{D}{d\sigma}\frac{d\tilde\rho^{(k)}}{d\sigma}}_{=0}, \left.\frac{d\tilde\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle\right\}\mathrm{d}\sigma + \left\langle \mathrm{grad}\, E|_{\rho^{(k)}}, \left.\frac{d\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle,\\
&= \frac{1}{\Delta t}\left\langle\left.\frac{d\tilde\rho^{(k)}}{d\sigma}\right|_{\sigma=1}, \left.\frac{d\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle + \left\langle \mathrm{grad}\, E|_{\rho^{(k)}}, \left.\frac{d\rho_\varepsilon^{(k)}}{d\varepsilon}\right|_{\varepsilon=0}\right\rangle.
\end{aligned}$$

Puisque $\left.\dfrac{d\rho_\varepsilon^{(k)}}{d\sigma}\right|_{\varepsilon=1}$ décrit tout $T_\rho\mathcal{M}$, nous obtenons :

$$\frac{1}{\Delta t}\left.\frac{d\tilde\rho^{(k)}}{d\sigma}\right|_{\sigma=1} + \mathrm{grad}\, E|_{\rho^{(k)}} = 0, \tag{7.51}$$

qui est la première variation. Ainsi, dans un cadre Euclidien, la première variation de (7.50) coïncide avec la méthode d'Euler implicite pour (7.31). Considérons à présent l'interpolation

$$\rho(t) = \tilde\rho^{(k)}\left(\frac{t-(k-1)\Delta t}{\Delta t}\right), \text{ pour } t\in[(k-1)\Delta t, k\Delta t].$$

Ainsi, la courbe

$$\begin{array}{rcl}[0,\infty[& \to & \mathcal{M}\\ t & \mapsto & \rho(t)\end{array}$$

est continue et différentiable par morceaux avec

$$\rho(k\Delta t) = \rho^{(k)} \qquad \text{et} \qquad \frac{d\rho}{dt}(k\Delta t^-) = \frac{1}{\Delta t}\left.\frac{d\tilde\rho^{(k)}}{d\sigma}\right|_{\sigma=1}.$$

Finalement par (7.51), nous obtenons

$$\frac{d\rho}{dt}(k\Delta t^-) = -\mathrm{grad}\, E|_{\rho(k\Delta t)}.$$

On a donc bien construit une approximation de la solution de (7.31).

7.3.2 Quelques rappels : distance de Wasserstein et transport optimal

Dans ce paragraphe nous rappelons la définition et quelques résultats classiques en transport optimal et la distance de Wasserstein. Le lecteur intéressé pourra se reporter aux ouvrages de Villani [Vil03]. Un rapide résumé des propriétés de la distance de Wasserstein peut être lu dans [CT07].

Dans [Ott01], l'auteur montre que la métrique d induite sur $(\mathcal{M}, g)$ est la métrique de Wasserstein. Donnons à présent une définition précise de cette distance.

Notation 1 *Dans toute la suite, nous noterons :*

- ⋆ $\mathcal{P}(\mathbb{R}^d)$*, l'espace des mesures de probabilité sur* $\mathbb{R}^d$ *;*
- ⋆ $\mathcal{P}_2(\mathbb{R}^d)$*, le sous–ensemble des mesures de second moment fini ;*
- ⋆ $\mathcal{P}_2^{ac}(\mathbb{R}^d)$*, le sous–ensemble des mesures absolument continues par rapport à la mesure de Lebesgue.*

Définition 7.3.2 *Soient μ, $\nu \in \mathcal{P}(\mathbb{R}^d)$ et T une application mesurable de $\mathbb{R}^d$ dans $\mathbb{R}^d$. On dit que T* transporte μ sur ν*, et on note $\nu = T\#\mu$ si pour tout ensemble mesurable $B \subset \mathbb{R}^d$, $\nu(B) = \mu \circ T^{-1}(B)$.*

On dit aussi que ν est le push-forward *ou* la mesure image *de μ par T,* i.e.

$$\int_{\mathbb{R}^d} \zeta \circ T(x)\, \mathrm{d}\mu(x) = \int_{\mathbb{R}^d} \zeta(y)\, \mathrm{d}\nu(y), \quad \forall \zeta \in C_b^0(\mathbb{R}^d).$$

Définition 7.3.3 (Distance de Wasserstein–Problème de Monge) *La distance de Wasserstein entre deux mesures μ et ν est donnée par*

$$d^2(\mu, \nu) := \inf_{T\,:\,\nu = T\#\mu} \int_{\mathbb{R}^d} |x - T(x)|^2 \, \mathrm{d}\mu(x).$$

La distance de Wasserstein peut également se réécrire sous la forme d'un problème variationnel relaxé. On obtient alors la définition équivalente suivante.

Définition 7.3.4 (Distance de Wasserstein–Problème de Kantorovitch) *La distance de Wasserstein entre deux mesures μ et ν est donnée par*

$$d^2(\mu, \nu) := \inf_{\pi \in \Gamma} \int_{\mathbb{R}^d \times \mathbb{R}^d} \left|x - y\right|^2 \, \mathrm{d}\pi(x, y).$$

où Γ est l'ensemble des mesures Boréliennes positives sur $\mathbb{R}^d \times \mathbb{R}^d$ de marginales μ et ν

Le problème de Kantorovitch (Définition 7.3.4) est un problème linéaire tandis que le problème de Monge (Définition 7.3.3) est un problème non–linéaire. Pour cette raison il est plus facile de montrer que l'infimum est atteint, et il existe une application optimale de transport $\pi(x, y) \in \Gamma$. Le théorème de Brenier [Bre91] montre sous une condition supplémentaire sur les mesures, que le graphe de cette application optimale est bien une application de transport $(x, T(x))$. Ainsi les problèmes de Monge et de Kantorovitch coïncident.

Théorème 7.3.1 (Brenier'91) *Supposons μ absolument continue par rapport à la mesure de Lebesgue. Alors il existe une application mesurable T telle que $\nu = T\#\mu$ et $T = \nabla\varphi$ avec φ une fonction convexe. Dans ce cas, on a*

$$d^2(\mu, \nu) = \int_{\mathbb{R}^d} \left|x - \nabla\varphi(x)\right|^2 \, \mathrm{d}\mu(x).$$

Dans le cadre de l'application du théorème de Brenier, la mesure optimale est donnée par $\pi_0 = (\mathrm{id}_{\mathbb{R}^d} \otimes \nabla\varphi)\,\#\mu$.

7.3.3 Écriture du schéma pour le système de Patlak–Keller–Segel (7.3)

La fonctionnelle d'énergie libre $\mathcal{G}$ associée au système (7.3) (avant 'blow–up') est donnée par [BCC08]

$$t \mapsto \mathcal{G}[\rho](t) := \mathcal{S}[\rho](t) + \mathcal{W}[\rho](t),$$

avec $\mathcal{S}[\rho](t)$, l'entropie de Boltzmann

$$\mathcal{S}[\rho](t) := \int_{\mathbb{R}} \rho(t, x) \log \rho(t, x)\, \mathrm{d}x$$

et $\mathcal{W}[\rho](t)$ l'énergie d'interaction

$$\mathcal{W}[\rho](t) := \frac{\chi}{\pi} \int \int_{\mathbb{R}\times\mathbb{R}} \rho(t, x)\rho(t, y) \log|x - y|\, \mathrm{d}x\, \mathrm{d}y.$$

La discrétisation en temps, selon le schéma de JKO donne l'algorithme suivant. Considérons un pas de temps $\Delta t > 0$ et une condition initiale $\rho_0 \in \mathcal{P}_2^{ac}(\mathbb{R}^d)$. La suite considérée $(\rho^n)_{n\in\mathbb{N}}$ est définie par $\rho^0 = \rho_0$ et

$$\rho^{n+1} \in \arg\inf_{\rho \in \mathcal{K}} \left\{ \mathcal{G}[\rho] + \frac{1}{2\Delta t} d(\rho^n, \rho)^2 \right\}, \tag{7.52}$$

où $\mathcal{K} \subset \mathcal{P}_2^{ac}(\mathbb{R}^d)$ est muni d'une structure Riemannienne compatible avec la distance d de Wasserstein. Dans [BCC08], les auteurs montrent l'existence de minimiseurs à (7.52) dans le cas sous–critique $\chi < \chi_c$, ainsi que la convergence de ce schéma lorsque le pas de temps Δt tend vers zéro.

7.3.4 Cas particulier de la dimension 1

Dans le cas de la droite réelle, considérons μ et ν deux mesures absolument continues par rapport à la mesure de Lebesgue, de densités respectives f et g. Soit $\nabla\varphi$ le transport optimal de μ sur ν donné par le théorème de BRENIER (Théorème 7.3.1). En dimension plus grande que un, φ, le transport optimal, est solution de l'équation de Monge–Ampère :

$$f = g \circ \nabla\varphi \, \mathrm{Det}\, D^2\varphi,$$

qui ne peut être résolue en général de façon explicite. Cependant en dimension un, comme nous le détaillons le système est plus simple.

Soient F et G les fonctions de répartition des fonctions f et g définies par

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y = \mu(]-\infty, x]).$$

Puisque, les mesures sont positives, leurs fonctions de répartition sont croissantes et il est facile de définir leur pseudo–inverse :

$$V(z) = F^{-1}(z) := \inf\{x \,:\, F(x) \geq z\}.$$

D'après le Théorème 7.3.1 et la définition de la mesure image (Définition 7.3.2), nous obtenons :

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y = \int_{-\infty}^{\varphi'(x)} g(y)\,\mathrm{d}y = G \circ \varphi'(x).$$

Ainsi, le transport optimal peut être donné explicitement comme $\varphi' = G^{-1} \circ F$ et la distance de Wasserstein peut–être exprimée sous la forme suivante

$$d^2(\mu, \nu) = \int_0^1 \left|F^{-1}(m) - G^{-1}(m)\right|^2 \,\mathrm{d}m, \tag{7.53}$$

beaucoup plus maniable et facile d'utilisation. La transformation $\rho \mapsto F^{-1}$ transforme la distance de Wasserstein en la distance Euclidienne entre les fonctions inverse de répartition. Cela a pour conséquence d'aplatir la métrique et d'améliorer grandement la formulation des flots de gradients associés.

Dans [BCC08], les auteurs utilisent cette propriété pour construire un schéma d'Euler implicite en temps pour le système (7.3) unidimensionnel, schéma qui correspond au schéma JKO via la représentation (7.53) de la distance de Wasserstein. Cette approche du schéma de JKO est également préférée à l'implémentation directe de (7.52) difficile dans le cas du système de Patlak–Keller–Segel, dû au défaut de convexité de la fonctionnelle $\mathcal{G}$: il manque une propriété de contraction pour la distance de Wasserstein.

La fin de ce paragraphe est consacrée à la description de leur schéma numérique. Dans ce qui suit nous décrivons les travaux de BLANCHET, CALVEZ et CARILLO [BCC08], puis à la Section suivante nous détaillerons notre schéma, extension de leurs travaux, à la lumière de la description du système après 'blow–up' (7.27).

Soient F_n et F_{n+1} les fonctions de répartition associée respectivement à ρ^n et ρ^{n+1}. D'après l'expression de la distance de Wasserstein sur la droite réelle (7.53), le schéma (7.52) peut être réécrit en $V_n = F_n^{-1}$ et $V_{n+1} = F_{n+1}^{-1}$, comme le flot de gradient de l'inverse de la fonction de répartition sous la structure associée à la distance L^2, soit :

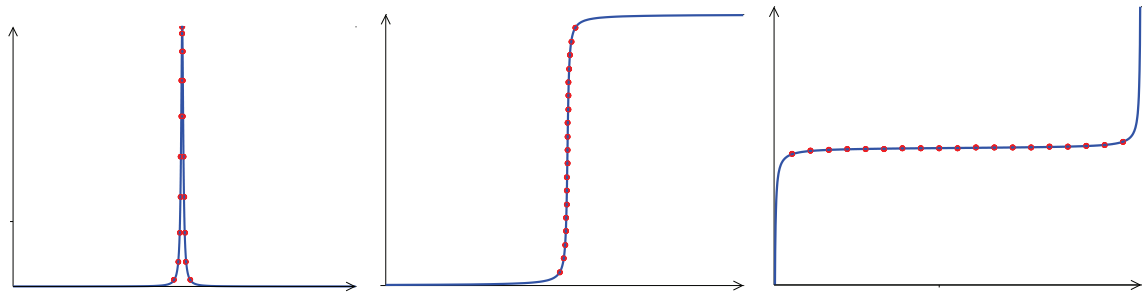
$$V_{n+1} \in \arg \inf_{\{W\,:\,(W^{-1})' \in \mathcal{K}\}} \left[\tilde{\mathcal{G}}[W] + \frac{1}{2\,\Delta t}\|W - V_n\|^2_{L^2(0,1)}\right],$$

où

$$\tilde{\mathcal{G}}[W] := -\int_0^1 \log W'(m)\,\mathrm{d}m + \frac{\chi}{\pi}\int_0^1\int_0^1 \log|W(m) - W(z)|\,\mathrm{d}m\,\mathrm{d}z.$$

Dans ce cas, la métrique est Euclidienne, et donc l'équation d'Euler–Lagrange associée à ce problème de minimisation est la suivante :

$$\frac{V_{n+1}(m) - V_n(m)}{\Delta t} = -\nabla\tilde{\mathcal{G}}[V_{n+1}],$$



(a) Fonction densité (b) Fonction de répartition (c) Inverse de la fonction de répartition

FIG. 7.3: Visualisation de la fonction densité, de sa fonction de répartition et de sa fonction de répartition inverse. Nous remarquons qu'un maillage uniforme pour l'inverse de la fonction de répartition entraine un raffinement automatique autour des points d'explosion.

où ∇ est l'opérateur gradient usuel sur $L^2(\mathbb{R})$. Cette équation d'Euler–Lagrange peut-être réécrite comme suit

$$-\frac{V_{n+1}(m)-V_n(m)}{\Delta t}=\frac{\partial}{\partial m}\left[\left(\frac{\partial V_{n+1}(m)}{\partial m}\right)^{-1}\right]+\chi H[V_{n+1}], \tag{7.54}$$

où H est lié à la transformée de Hilbert et défini par

$$H[V](m):=\frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|V(m)-V(z)|\geq\varepsilon}\frac{1}{V(m)-V(z)}\,\mathrm{d}z.$$

Reste à discrétiser également en espace. La grille est uniforme, notons Δm le pas d'espace, constant. Si nous posons $V_n^i := V_n(i\,\Delta m)$, pour tout $i=0,\dots,I$, et $I\,\Delta m=1$, la discrétisation en espace par différences finies de (7.54) est la suivante :

$$-\frac{V_{n+1}^i-V_n^i}{\Delta t}=\frac{1}{V_{n+1}^{i+1}-V_{n+1}^i}-\frac{1}{V_{n+1}^i-V_{n+1}^{i-1}}+\frac{\chi\,\Delta m}{\pi}\lim_{\varepsilon\to 0}\sum_{j:|V_{n+1}^i-V_{n+1}^j|\geq\varepsilon}\frac{1}{V_{n+1}^i-V_{n+1}^j},$$

avec la condition initiale $V_0=\rho_0$. Nous imposons des conditions de Neumann aux bords pour le problème exprimé en la fonction ρ ce qui nous donne après changement de fonction, pour tout n,

$$\frac{1}{V_n^I-V_n^{I-1}}=0\quad\text{et}\quad\frac{1}{V_n^1-V_n^0}=0,$$

de sorte que le centre de masse est conservé :

$$\forall n,\quad\sum_{i=0}^{I}V_n^i=0.$$

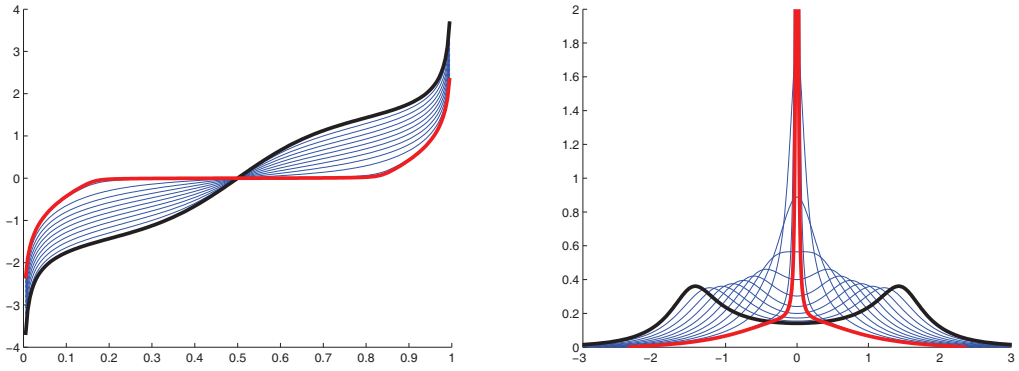
À chaque pas de temps, la solution de ce système non–linéaire est obtenue par une procédure de Newton–Raphson.

Remarque 19 *Nous visualisons à la Figure 7.3, le changement de fonction en dessinant le graphe de chacune de ces fonctions. L'essentiel à retenir pour la suite est qu'à un pic pour la densité correspond un plateau pour l'inverse de la fonction de répartition.*

Nous donnons les résultats numériques pour ce schéma à la Figure 7.4. Pour cette simulation, nous avons choisi $\chi=5/2$ et la condition initiale

$$V_0(m)=\frac{2m-1}{(m(1-m))^{1/4}}.$$

Nous présentons les résultats en terme de fonction de répartition inverse Figure 7.4(a) et pour la densité Figure 7.4(b), aux mêmes temps. Nous remarquons, au fur et à mesure que la densité forme un pic, l'établissement d'un plateau sur le graphe de la fonction de répartition inverse.



(a) (b)

FIG. 7.4: Résultats numériques pour le schéma avant explosion, en trait noir la condition initiale, en trait rouge la condition à l'arrêt de l'algorithme ; $\chi = 5/2$ et $V_0(m) = \dfrac{e^{10(m-0.5)} - 1}{[m(1-m)]^{1/4}}$.

7.4 Adaptation du schéma après blow–up

Dans [BCC08], comme nous l'avons décrit précédemment, les auteurs donnent un schéma numérique pour le problème (7.3) en dimension $d = 1$, pour des temps avant explosion. Pour ce faire, ils écrivent le problème en termes de distance de Wasserstein et ils obtiennent des équations pour l'inverse de la fonction de répartition de la densité ρ.

Notre but est, en s'inspirant leurs idées, d'étendre ce schéma aux temps après explosion en s'appuyant sur la description du système donnée par le système (7.27).

L'idée de base est la suivante : puisqu'une masse de Dirac pour la densité correspond à un plateau en terme de fonction de répartition inverse, nous suivons la décomposition de la solution en partie continue et masses de Dirac (7.26) : nous "coupons" les parties aplaties pour mettre en évidence les masses de Dirac, puis nous "recollons" les deux branches restantes qui correspondent à la partie continue.

Soyons un peu plus précis et écrivons l'équation (7.27) en termes de fonction de répartition inverse.

Soit la fonction de répartition $\bar{F}$ correspondant à la partie continue,

$$\bar{F}(x) = \int_{-\infty}^{x} \bar{\rho}(m)\,\mathrm{d}m, \tag{7.55}$$

et $\bar{V} = \bar{F}^{-1}$, son pseudo–inverse.

L'équation vérifiée par la fonction $\bar{V}$ est la suivante :

$$\partial_t \bar{V} = \frac{\partial}{\partial m}\left[\left(\frac{\partial \bar{V}(m)}{\partial m}\right)^{-1}\right] + \chi H_1[\bar{V}] + \chi H_2[\bar{V}, \hat{V}], \tag{7.56}$$

où H_1 est le terme correspondant aux interactions entre les particules de la partie continue :

$$H_1[\bar{V}] = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|\bar{V}(m)-\bar{V}(z)|\geq\varepsilon} \frac{1}{\bar{V}(m) - \bar{V}(z)}\,\mathrm{d}z,$$

et H_2 le terme correspondant aux interactions entre la partie continue et les masses de Dirac :

$$H_2[\bar{V}, \hat{V}] = \frac{1}{\pi}\sum_{\ell=1}^{L} \frac{M^\ell}{\bar{V}(m) - x^\ell},$$

où M^ℓ et x^ℓ sont respectivement la masse et la position de la ℓ^e masse de Dirac de $\hat{V}$. Notons que s'il n'y a pas de masse de Dirac, ce terme s'annule et nous retrouvons l'équation (7.54).

Nous omettons dans la suite la barre, étant convenu que la notation V correspond à la partie continue, la partie discrète étant décrite par les masses M^ℓ et les positions x^ℓ des masses de Dirac.

Nous notons V_n^i l'approximation de la solution au temps $t_n = n\,\Delta t$ au point $w = i\Delta m$, $i = 1, \dots, I$, $I\Delta m = 1$. Nous discrétisons le problème comme suit :

$$-\frac{V_{n+1}^i - V_n^i}{\Delta t} = \frac{1}{V_{n+1}^{i+1} - V_{n+1}^i} - \frac{1}{V_{n+1}^i - V_{n+1}^{i-1}} + \chi\left(\frac{\Delta m}{\pi} \sum_{|V_{n+1}^i - V_{n+1}^j| \geq \varepsilon} \frac{1}{V_{n+1}^i - V_{n+1}^j} + \frac{1}{\pi}\sum_{\ell=0}^{L} \frac{M^\ell}{V_{n+1}^i - x^\ell}\right), \quad (7.57)$$

avec un pas d'espace constant égal à Δm. La méthode des différences finies que nous utilisons est un schéma d'Euler semi–implicite. Une procédure de Newton–Raphson est utilisée pour résoudre ce schéma implicite. Nous imposons des conditions de Neumann pour le problème posé pour la densité (7.27), *i.e.* pour tout n, $\dfrac{1}{V_n^I - V_n^{I-1}} = 0$ et $\dfrac{1}{V_n^1 - V_n^0} = 0$, de sorte que le centre de masse soit conservé.

Remarque 20 *En pratique, pour gagner en rapidité, nous utilisons une discrétisation semi–implicite et l'équation* (7.57) *est remplacée l'équation discrète suivante :*

$$-\frac{V_{n+1}^i - V_n^i}{\Delta t} = \frac{1}{V_{n+1}^{i+1} - V_{n+1}^i} - \frac{1}{V_{n+1}^i - V_{n+1}^{i-1}} + \chi\left(\frac{\Delta m}{\pi} \sum_{|j-i|\leq 1, |V_{n+1}^i - V_{n+1}^j| \geq \varepsilon} \frac{1}{V_{n+1}^i - V_{n+1}^j} + \frac{\Delta m}{\pi} \sum_{|j-i|>1, |V_{n+1}^i - V_n^j| \geq \varepsilon} \frac{1}{V_{n+1}^i - V_n^j} + \frac{1}{\pi}\sum_{\ell=0}^{L} \frac{M^\ell}{V_{n+1}^i - x^\ell}\right), \quad (7.58)$$

Nous réalisons ainsi des simulations avec une condition initiale V_0 choisissant le paramètre χ de sorte qu'il apparaisse des points de concentration en temps fini. L'explosion correspond à l'apparition d'un plateau dans le graphe de l'inverse de la fonction de répartition. Toute la difficulté consiste à déterminer numériquement quel est le temps d'explosion et quelle partie du plateau constitue la masse de Dirac, *i.e.* combien nous devons couper.

7.4.1 Heuristique

Une manière de répondre à cette question est de reprendre l'algorithme de [BCC08] décrit à la Section 7.3.4 et de comprendre pourquoi le schéma échoue à l'approche de l'explosion.

Cet algorithme s'arrête juste avant l'explosion, à chaque fois par défaut de convergence de la procédure de Newton–Raphson. Pourquoi, à l'explosion, la procédure de Newton ne converge-t-elle pas ? L'idée est basée sur la remarque suivante. Supposons que pour des raisons de symétrie de la condition initiale, on sache que la masse de Dirac apparaîtra en $x = 0$. A l'approche de l'explosion, la procédure de Newton–Raphson doit résoudre une équation similaire à

$$V_{n+1}^i + \frac{\Delta t}{V_{n+1}^i} = V_n^i,$$

qui n'a manifestement pas de solution si V_n est trop proche de 0 (à distance de l'ordre de $\sqrt{\Delta t}$). Dans ce cas, on doit considérer que V_{n+1}^i est en 0 et appartient au plateau, c'est–à–dire à la masse de Dirac.

7.4.2 Algorithme « cut & paste »

Avant de rentrer dans les détails de l'algorithme, nous voulons ici en donner les principes de base.

Suite à la remarque précédente, Section 7.4.1, nous considérons que lorsque la procédure de Newton–Raphson ne converge pas, nous devons prélever de la masse dans la partie continue pour la replacer dans un ou plusieurs Dirac. Il y a plusieurs façons de procéder : nous pouvons soit créer une nouvelle masse de Dirac, soit augmenter la masse d'un des Dirac existants. Une masse de Dirac, c'est–à–dire un plateau de la fonction de répartition inverse, est repérée par le point de plus faible pente. La longueur des plateaux est déterminée ainsi : nous nous donnons un seuil η (le choix de η est discuté au paragraphe suivant) et tout point dont la distance à une masse de Dirac est

inférieure à η est considéré comme appartenant au plateau correspondant et donc incorporé à la masse de Dirac. Lorsqu'un plateau est ainsi repéré, nous enlevons les points correspondants de la partie continue et les remplaçons par un unique point ayant pour ordonnée la position de la masse de Dirac correspondante. La masse du Dirac est mise à jour avec la longueur du plateau prélevé qui correspond exactement à l'ajout de masse. Notons qu'ainsi le nombre de points de la discrétisation, noté I, de la partie continue diminue. L'algorithme s'arrête lorsqu'il n'en reste que 5.

Enfin, nous ne devons pas oublier de déplacer les masses de Dirac selon l'équation (7.27). Si deux masses de Dirac se trouvent trop proches, à une distance de l'ordre de η, elles fusionnent.

Dans le paragraphe qui suit, nous détaillons pas à pas les étapes de notre algorithme. L correspond aux nombres de masses de Dirac, I le nombre de points de discrétisation pour la partie continue. M^ℓ (resp. x^ℓ) est la masse (resp. la position) de la ℓ^e masse de Dirac. `flag` est un drapeau qui indique le nombre de passage dans la procédure « Cut & Paste » ; au premier passage, nous regardons si la taille des plateaux déjà existants a augmenté. Au deuxième passage, qui signifie que la procédure précédente n'a pas suffi, nous regardons si d'autres masses de Dirac sont apparues.

Algorithme. Plaçons nous au temps $T = t^n = n\,\Delta t$ et supposons que nous avons calculé les V_k^i pour $k \leq n$ et $i \in [\![0, I]\!]$.
`flag = 1`

`I ≥ 5`
η fixé

Étape 1 Calcul de V_{n+1}^i pour $i = 0, \ldots, I$ solution de l'équation discrète (7.57) à l'aide d'une procédure de Newton–Raphson.

- ⋆ l'algorithme de Newton–Raphson converge.
  `Aller à l'`**Étape 2**`.`
  `flag ⟵ 1`
- ⋆ l'algorithme de Newton–Raphson ne converge pas.
  `Réaliser la procédure « Cut & Paste ». Aller à l'`**Étape CP**`.`

Étape 2 Déplacer les masses de Dirac selon l'Équation (7.27) discrétisée de la façon la plus simple :

$$x_{n+1}^\ell = x_n^\ell - \Delta t \frac{\chi}{\pi} \left(\sum_{|x_n^\ell - V_n^i| \geq \varepsilon} \frac{\Delta m}{x_n^\ell - V_n^i} + \sum_{k \neq \ell} \frac{M_n^k}{x_n^\ell - x_n^k} \right). \tag{7.59}$$

`Aller à l'`**Étape 3**`.`

Étape 3 **Procédure de fusion des masses de Dirac**

- ⋆ Si pour tout $j \neq k$, $|x^j - x^k| > 2\eta$,
  `Aller à l'`**Étape 4**`.`
- ⋆ Sinon, si $|x^j - x^k| \leq 2\eta$, nous devons fusionner les deux masses de Dirac correspondantes.
  - Calcul de la position de la masse de Dirac obtenue par fusion ;
    $x^j \longleftarrow \left(M^j x^j + M^k x^k\right) / (M^j + M^k)$;
  - Calcul de la nouvelle masse ;
    $M^j \longleftarrow M^j + M^k + |j - k|\,\Delta m$;
  - Couper la partie continue située éventuellement entre les deux masses de Dirac. Voir Figure 7.5.
  - `Aller à l'`**Étape 4**`.`

Étape 4 Faire avancer le temps.
`T ⟵ T +` Δt
`n ⟵ n+1`
`Aller à l'`**Étape 1**`.`

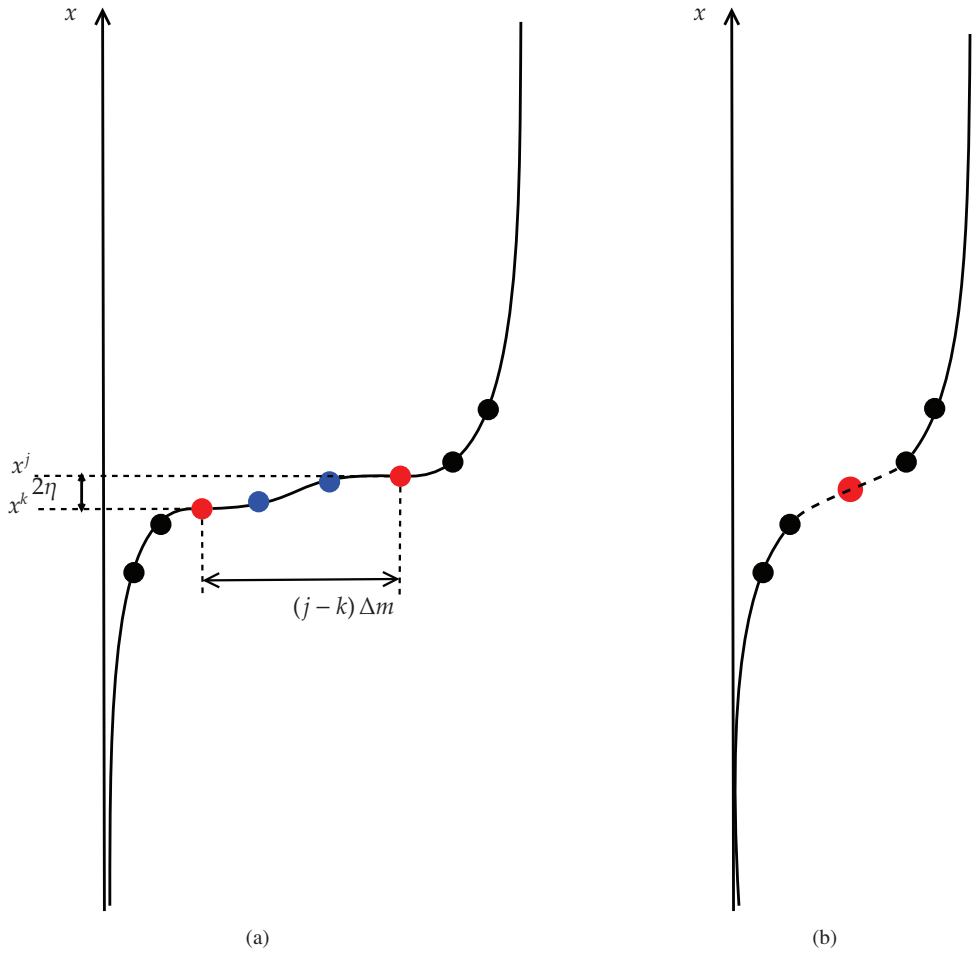


FIG. 7.5: Fusion de deux masses de Dirac

Étape CP **Procédure « Cut & Paste ».**

★ Si `flag=1` :

Étape CP– 1 (a) Si `L` $\geq$ `1`,
`Aller à 1'`**Étape CP– 2**.

(b) Sinon `L=0`,

- `L` $\longleftarrow$ `1` ;
- Déterminer le point de pente minimale s_{min}, (i, V_n^i) ;
- Stocker cette pente comme pente de référence.
  $s_R \longleftarrow s_{min}$;
- Mémoriser l'emplacement de la nouvelle masse de Dirac ;
  $x^1 \longleftarrow V_n^i$. Voir Figure 7.6.
- `Aller à 1'`**Étape CP– 2**.

Étape CP– 2 Tout point à une distance au plus η d'une masse de Dirac est considéré comme appartenant à la masse de Dirac.

- Repérer tous les points à une distance η de la ℓ^e masse de Dirac située en x_n^ℓ. C'est un plateau, on repère l'indice de début i_{deb} et de fin i_{fin} du plateau. Voir Figure 7.7(a).
- Calculer la masse correspondant à cette partie ;
  `m` $\longleftarrow (i_{fin} - i_{deb})\Delta m$;

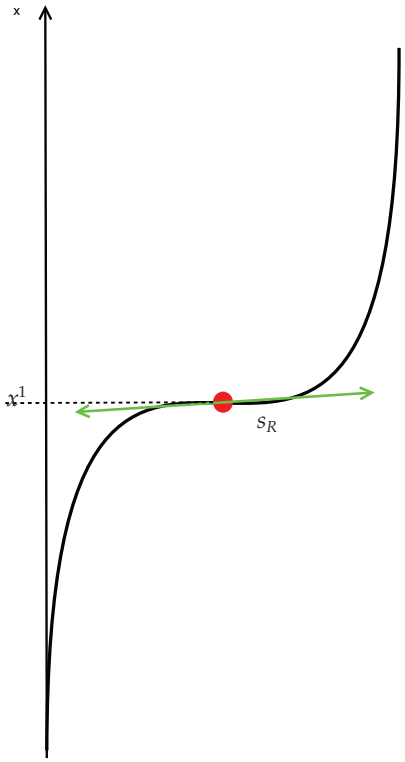


FIG. 7.6: Repérage de la première masse de Dirac

- On remplace ce plateau par un seul point d'ordonnée x_n^ℓ ;
- Mettre à jour la masse du ℓ^e Dirac ;
  $M^\ell \longleftarrow M^\ell$ `+ m;`
- Recoller les deux branches (Figure 7.7(b) ;
  `I` $\longleftarrow$ `I -(`$i_{fin} - i_{deb}$`);`
- `flag` $\longleftarrow$ `0;`
- `Aller à l'`**Étape 1**.

★ Si `flag=0` : apparition d'une nouvelle masse de Dirac.

Étape CP– ★
- Repérer le point de pente minimale s_{min} ; (i, V_n^i). Si $s_{min} = s_R \pm 10\%$, nous décidons que nous avons une nouvelle masse de Dirac en ce point.
- $X^{L+1} \longleftarrow V_n^i$;
- `L` $\longleftarrow$ `L+1;`
- `flag` $\longleftarrow$ `1;`
- `Aller à l'`**Étape 1**.

Remarque 21 (À propos du maillage) *Un maillage régulier pour l'inverse de la fonction de répartition semble approprié au problème que nous regardons. En effet, il concentre les points de discrétisation autour de plateaux pour l'inverse de la fonction de répartition ce qui se répercute par une concentration des points de discrétisation autour des pics de la densité. Nous avons illustré cela à la Figure 7.3, en matérialisant par des points rouges, les points de discrétisation correspondant à un maillage régulier pour l'inverse de la fonction de répartition.*

Le fait de perdre des points au fur et à mesure de l'algorithme pour la description de la partie continue n'entraine pas de perte de précision, car chaque maille correspond au même élément de masse Δm, quel que soit le nombre de mailles.

Remarque 22 (À propos de la fusion de deux masses de Dirac) *En pratique, lorsqu'on fusionne deux masses de Dirac, les plateaux* (i.e. *les deux points rouges sur la Figure 7.7(a)) ne sont distants que d'une maille.*

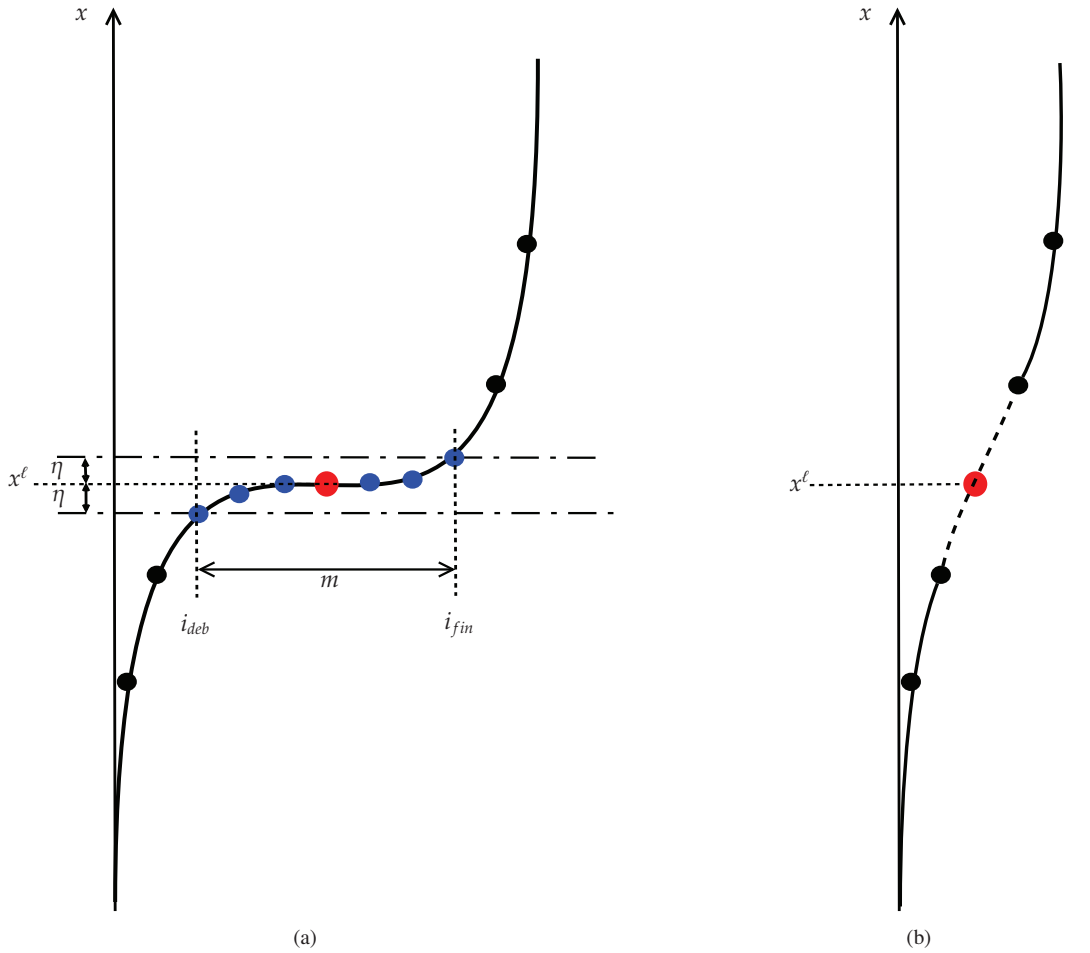


FIG. 7.7: Repérage d'un plateau, découpe et recollage

Pour la fusion de deux masses de Dirac, nous choisissons de placer la masse de Dirac résultante de cette fusion au centre de gravité du système formé par les masses de Dirac. Ce choix correspond à négliger toute la masse en dehors de ces deux points d'accumulation ; dans ce cas, d'après la théorie du problème à deux corps le système conserve son centre de masse et la masse de Dirac résultante se situe logiquement au centre de gravité du système.

Remarque 23 *Le système* (7.57) *ainsi que l'algorithme décrit ci–dessus s'apparentent à un schéma de particules pour l'équation de Keller–Segel* (7.1) *[HS09]. En effet V décrit en quelque sorte la position du nuage de particules. Une différence majeure néanmoins réside dans le fait que les particules restent ordonnées au cours du temps (condition de croissance de V). La diffusion Brownienne est remplacée par un opérateur déterministe équivalent au Laplacien.*

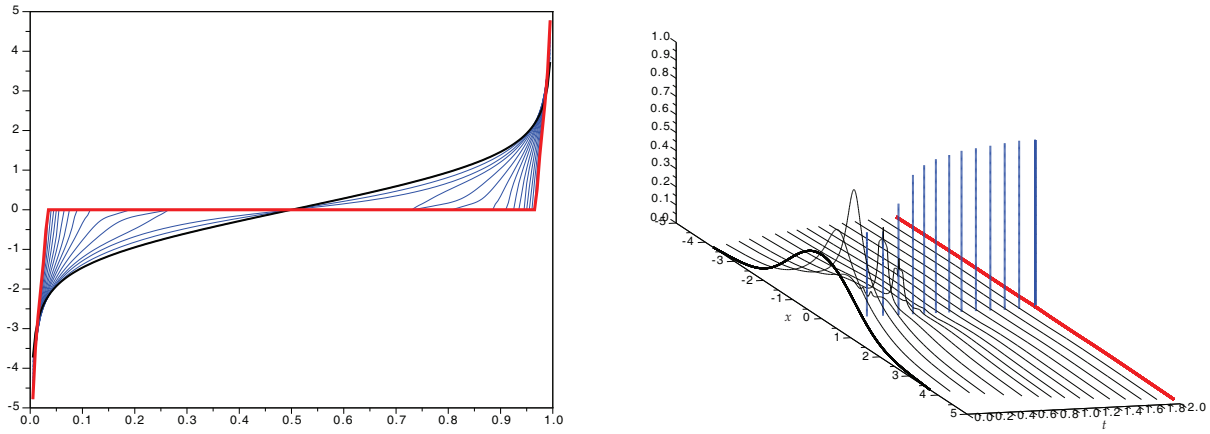
Une condition CFL ou comment choisir η ?

Une question reste encore en suspend : comment choisir le seuil η ? Nous avons déjà mentionné certains indices indiquant que η devrait être de l'ordre de $\sqrt{\Delta t}$.

Dans le cas d'une discrétisation explicite, la condition CFL donnée dans [GT06] se traduit dans notre cas par :

$$\eta \leq \sqrt{6\Delta t}. \tag{7.60}$$

Il semble qu'il existe également une telle condition pour un schéma implicite. Pour notre algorithme on fixe $\eta = \sqrt{6\Delta t}$ la condition CFL.



(a) (b)

FIG. 7.8: Évolution de la solution dans les deux variables dans le cas symétrique avec une seule masse de Dirac

Remarque 24 *Dans [HS09], les auteurs donnent une description stochastique des particules pour le système de Keller–Segel 2D. Ils utilisent la théorie des deux corps pour décrire le mouvement de deux particules lourdes assez proches et trouvent que deux telles particules entrent en collision si et seulement si la distance entre elles d vérifie :*

$$d \leq \sqrt{\frac{M^1 + M^2}{\pi} \Delta t}.$$

Le lecteur notera que cette condition est comparable à celle que nous choisissons dans notre schéma.

7.4.3 Illustrations numériques

Nous réalisons des simulations numériques dans différents cas. Dans tous les cas, le pas de temps choisi est $\Delta t = 5 \cdot 10^{-3}$.

Cas symétrique, une seule masse de Dirac Nous choisissons $\chi = 5/2\pi$ et la condition initiale symétrique,

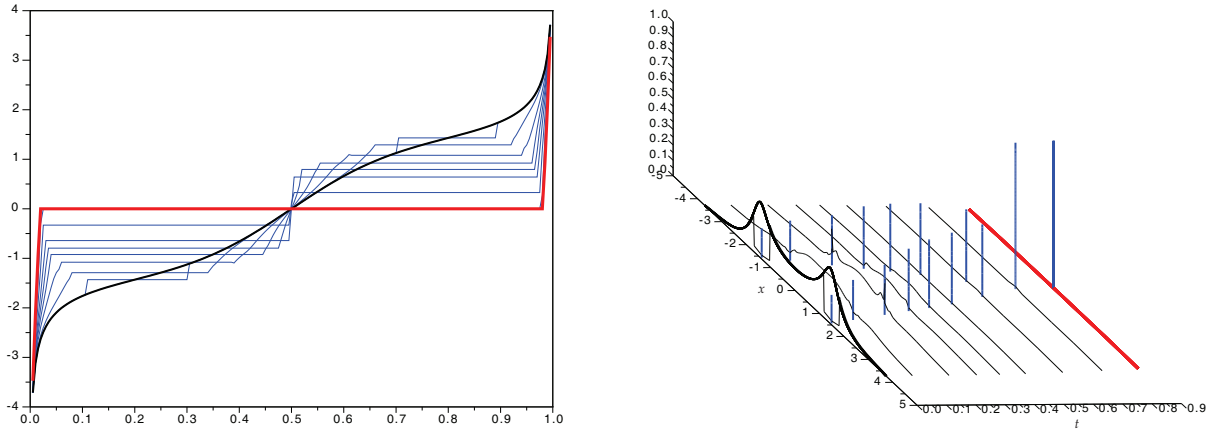
$$V_0(m) = \frac{2m - 1}{(m(1-m))^{1/4}}.$$

En posant $\chi = 5/2\pi$, nous sommes certains de n'obtenir qu'une seule masse de Dirac car $2\pi \leq \chi < 4\pi$. Nous présentons à la Figure 7.8 les résultats numériques sous deux formes. Nous donnons à la Figure 7.8(a) l'évolution de la fonction de l'inverse de la fonction de répartition, que l'on a reconstruit avec les plateaux. En trait gras on donne la condition initiale, en trait rouge la condition finale. La Figure 7.8(b) représente la fonction densité correspondante avec la masses de Dirac représentées par les barres dont la longueur représente la masse du Dirac.

Cas symétrique, deux masses de Dirac Nous choisissons cette fois $\chi = 5\pi$ et une condition initiale toujours symétrique,

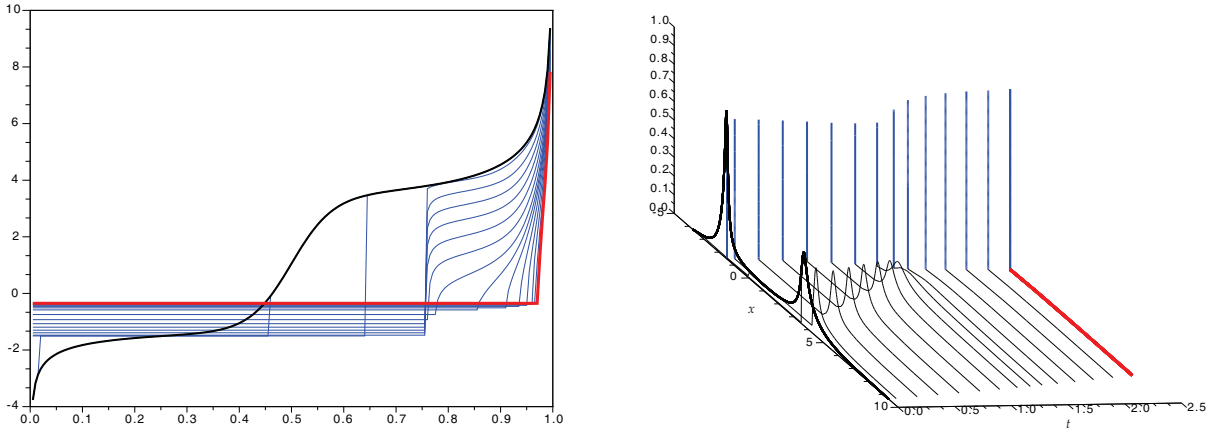
$$V_0(m) = \frac{-1 + 2e^{10m-5}}{1 + 2e^{10m-5}} \frac{1}{(m(1-m))^{1/4}}.$$

Nous présentons à la Figure 7.9 les résultats numériques sous deux formes. Nous donnons à la Figure 7.9(a), l'évolution de la fonction de l'inverse de la fonction de répartition, que l'on a reconstruit avec les plateaux. Les conventions pour la visualisation sont les mêmes que dans le cas précédent. La Figure 7.9(b) représente la fonction densité correspondante.



(a) (b)

FIG. 7.9: Évolution de la solution dans les deux variables dans le cas symétrique avec deux masses de Dirac



(a) (b)

FIG. 7.10: Évolution de la solution dans les deux variables dans le cas non symétrique pour une condition initiale présentant deux pics d'amplitudes inégales

Cas non symétrique Nous choisissons cette fois $\chi = 10\pi$ et une condition initiale cette fois non symétrique, présentant deux pics d'amplitudes inégales :

$$V_0(m) = \frac{-1 + 2e^{25(m-0.45)}}{1 + 2e^{25(m-0.5)}} \frac{1}{(m(1-m))^{1/4}}.$$

Nous présentons à la Figure 7.10 les résultats numériques sous deux formes. Nous donnons à la Figure 7.10(a), l'évolution de la fonction de l'inverse de la fonction de répartition, que l'on a reconstruit avec les plateaux. Les conventions pour la visualisation sont les mêmes que dans les cas précédents. La Figure 7.10(b) représente la fonction densité correspondante. Nous observons que dans ce cas une seule masse de Dirac apparaît, l'autre 'bosse' de concentration se faisant absorbée peu à peu vers le point de concentration.

Remarque 25 *Au Lemme 1, nous avons conjecturé que la partie régulière de la densité s'annule aux points de discontinuité. Cette conjecture se confirme numériquement. La Figure 7.11 représente la densité de cellules et la masse de Dirac pour une condition initiale de type gaussienne, symétrique et* $\chi = 10\pi$. *On observe bien l'annulation de la partie continue à l'emplacement de la discontinuité, ce qui est en faveur de la conjecture.*

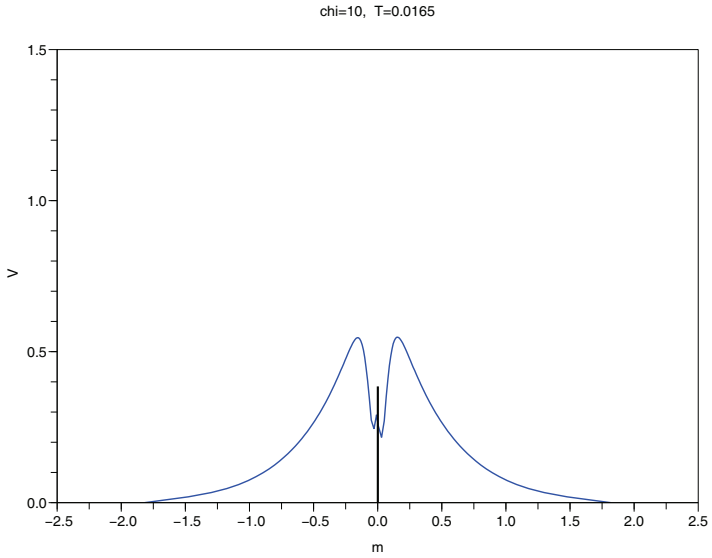


FIG. 7.11: Illustration de la conjecture du Lemme 1

7.5 Conclusion et perspectives

Nous avons pu décrire, à l'aide de la notion de mesure de défaut, le système de Patlak–Keller–Segel modifié en dimension $d = 1$ après explosion. La difficulté par rapport aux travaux de DOLBEAULT et SCHMEISER [DS09] réside essentiellement dans l'apparition en dimension $d = 1$ de la transformée de Hilbert tandis qu'en dimension $d = 2$ le noyau logarithmique est la solution fondamentale de l'équation de Poisson. La formulation obtenue nous a permis d'étendre le schéma de BLANCHET *et al.* [BCC08] aux temps après explosion. Bien que ce schéma soit fondé sur des notions fines liées à l'utilisation d'une nouvelle structure Riemannienne, il s'exprime facilement en dimension $d = 1$ car la distance de Wasserstein a une expression explicite dans ce cas.

Il existe de nombreuses perspectives à ce travail. Outre la démonstration du Lemme 1 qui est un travail encore en cours, nous avons remarqué que numériquement, lorsque la masse de Dirac apparaît, la longueur du plateau correspondant, c'est–à–dire la masse concentrée au point d'accumulation, est exactement égale à la masse critique, à une taille de maille près. Une analyse mathématique plus approfondie pourrait peut-être confirmer cette observation.

À plus long terme, nous pourrions envisager d'étendre ce type de schéma aux dimensions supérieures. Un travail allant dans ce sens a été réalisé par CARILLO et MOLL. Les auteurs proposent dans [CM10] un schéma numérique pour des équations non–linéaires écrites en coordonnées Lagrangiennes. À nouveau leur travail est lié à l'étude d'équations de flots de gradients par rapport à la distance de Wasserstein.

Remerciements

Je tiens à remercier le CRM (Centre de Recerca Matemàtica) de Barcelone pour l'accueil et l'ambiance propice au travail et aux bonnes idées. Un grand merci à Antoine et Vincent pour leur lecture attentive et minutieuse de ce chapitre et leurs remarques.

Bibliographie

[BCC08] Adrien Blanchet, Vincent Calvez, and Jos´ A. Carrillo.
Convergence of the mass–transport steepest descent scheme for the subcritical Patlak–Keller–Segel model.
SIAM J. Numer. Anal., 46(2) :691–721, 2008.

[BDP06] BlancheAdrien, Jean Dolbeault, and Benoît Perthame.
Two–dimensional Keller-Segel model : optimal critical mass and qualittive properties of the solutions.
Electron. J. Differential Equations, (44), 2006.

[Bil09] Frédérique Billy.
Modélisation mathématique multi-échelle de l'angiogenèse tumorale. Analyse de la réponse tumorale aux traitements anti-angiogéniques.
Thèse de doctorat, Université Claude Bernard – Lyon I, décembre 2009.

[BP00] Nicola Bellomo and Luigi Preziosi.
Modelling and mathematical problems related to tumor evolution and its interaction with the immune system.
Math. Comput. Modelling, 32(3-4) :413–452, 2000.

[Bre91] Yann Brenier.
Polar factorization and monotone rearrangement of vector-valued functions.
Comm. Pure Appl. Math., 44(4) :375–417, 1991.

[CM10] José A. Carrillo and J. Salvador Moll.
Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms.
SIAM J. Sci. Comput., 31(6) :4305–4329, 2009/10.

[CPSt] Vincent Calvez, Benoît Perthame, and Mohsen Sharifi tabar.
Modified Keller–Segel system and critical mass for the log interaction kernel.
Contemp. Math., 429 :45–62.

[CPZ04] Lucilla Corrias, Benoît Perthame, and Hatem Zaag.
Global solutions of some chemotaxis and angiogenesis systems in high space dimensions.
Milan J. Math., 72 :1–28, 2004.

[CT07] José A. Carrillo and Giuseppe Toscani.
Contractive probability metrics and asymptotic behavior of dissipative kinetic equations.
Riv. Mat. Univ. Parma (7), 6 :75–198, 2007.

[DP04] Jean Dolbeault and Benoît Perthame.
Optimal critical mass in the two-dimensional Keller-Segel model in $\mathbb{R}^2$.
C. R. Math. Acad. Sci. Paris, 339(9) :611–616, 2004.

[DS09] Jean Dolbeault and Christian Schmeiser.
The two-dimensional Keller-Segel model after blow-up.
Discrete Contin. Dyn. Syst., 25(1) :109–121, 2009.

[GT06] Laurent Gosse and Giuseppe Toscani.
Lagrangian numerical approximations to one-dimensional convolution-diffusion equations.
SIAM J. Sci. Comput., 28(4) :1203–1227 (electronic), 2006.

[HP04] Thomas Hillen and Alex Potapov.

The one-dimensional chemotaxis model : global existence and asymptotic profile.
Math. Methods Appl. Sci., 27(15) :1783–1801, 2004.

[HPS07] Thomas Hillen, Kevin Painter, and Christian Schmeiser.
Global existence for chemotaxis with finite sampling radius.
Discrete Contin. Dyn. Syst. Ser. B, 7(1) :125–144 (electronic), 2007.

[HS09] Jan Haškovec and Christian Schmeiser.
Stochastic particle approximation for measure valued solutions of the 2d Keller-Segel system.
J Stat Phys, (135) :133–151, 2009.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto.
The variational formulation of the Fokker-Planck equation.
SIAM J. Math. Anal., 29(1) :1–17, 1998.

[JL92] Willi Jäger and Stephan Luckhaus.
On explosions of solutions to a system of partial differential equations modelling chemotaxis.
Trans. Amer. Math. Soc., 329(2) :819–824, 1992.

[Ott01] Felix Otto.
The geometry of dissipative evolution equations : the porous medium equation.
Comm. Partial Differential Equations, 26(1-2) :101–174, 2001.

[Pou02] Frédéric Poupaud.
Diagonal defect measures, adhesion dynamics and Euler equation.
Methods Appl. Anal., 9(4) :533–561, 2002.

[Sch81] Laurent Schwartz.
Cours d'analyse. 1.
Hermann, Paris, second edition, 1981.

[Vel04a] Juan J. L. Velázquez.
Point dynamics in a singular limit of the Keller-Segel model. I. Motion of the concentration regions.
SIAM J. Appl. Math., 64(4) :1198–1223 (electronic), 2004.

[Vel04b] Juan J. L. Velázquez.
Point dynamics in a singular limit of the Keller-Segel model. II. Formation of the concentration regions.
SIAM J. Appl. Math., 64(4) :1224–1248 (electronic), 2004.

[Vel04c] Juan J. L. Velázquez.
Well-posedness of a model of point dynamics for a limit of the Keller-Segel system.
J. Differential Equations, 206(2) :315–352, 2004.

[Vil03] Cédric Villani.
Topics in optimal transportation, volume 58 of *Graduate Studies in Mathematics*.
American Mathematical Society, Providence, RI, 2003.

[Wei04] Cornelis J. Weijer.
Dictyostelium morphogenesis.
Curr. Opin. Genet. Dev., 2004.

[Wei09] Cornelis J. Weijer.
Collective cell migration in development.
(122(Pt 18)) :3215–23, September 2009.

Résumé

Modélisation, analyse mathématique et simulation numérique de problèmes issus de la biologie

Cette thèse est consacrée à l'étude de quatre problèmes issus de la biologie.

Le premier concerne la modélisation d'une population de métastases. Le modèle abouti a une équation de McKendrick-Von Foerster : une équation de conservation munie d'un terme au bord non–local. Nous montrons l'existence d'une unique solution et étudions son comportement asymptotique à l'aide de la notion d'entropie relative généralisée. L'étude numérique utilise le schéma WENO.

Le deuxième concerne la modélisation de la respiration. Nous étudions la simulation des flux d'air dans l'appareil respiratoire à l'aide d'un modèle multi–échelle. Le système obtenu possède des conditions aux bords dissipatives non–usuelles. La méthode numérique employée est une méthode de décomposition qui permet de réduire le problème à la résolution de problèmes de Stokes avec conditions aux bords de type Dirichlet–Neumann classiques. Puis nous proposons un modèle pour les échanges gazeux montrant l'hétérogénéité de l'absorption de l'oxygène le long de l'arbre bronchique.

La troisième partie concerne la cascade MAPK dans des ovocytes de Xénopes. La modélisation amène à une équation de type KPP. Après une étude mathématique montrant l'existence d'un front d'onde, nous réalisons une étude numérique fine du système.

Enfin, nous étudions le système de Patlak–Keller–Segel 1D après explosion. Après une étude mathématique permettant de décrire le système après explosion à l'aide d'une mesure de défaut, nous donnons un schéma numérique adoptant le point de vue du transport optimal et permettant de simuler le système après explosion.

Mots-clés: biologie, métastases, McKendrick–Von Foerster, respiration, Navier–Stokes, conditions dissipatives, effets d'écrantage, MAP Kinase, front d'onde, Keller–Segel, mesure de défaut, transport optimal.

Abstract

Modelisation, mathematical analysis and numerical simulation of problems coming from biology.

We investigate four models coming from biological contexts.

The first one concerns a model describing the growth of a population of tumors. This model leads to a McKendrick–Von Foerster equation : a conservation law with a non–local boundary condition. We prove the existence and unicity of a solution, then we study, using the general relative entropy, its asymptotic behavior. We provide numerical simulations using WENO scheme.

The second part concerns the modelisation of the respiration. First we study the air flux in the bronchial tree using a mulstiscale model. The system present non–usual dissipative boundary conditions. The numerical scheme we use is based on a decomposition idea that reduce the system to the resolution of Stokes problems with standard Dirichlet–Neumann conditions. Then, we propose a model concerning the gas exchanges bringing to light the heterogeneity of the absorption of oxygen along the bronchial tree.

The third part concerns the MAPK cascade in Xenopus oocytes. The modelisation leads to an equation of KPP type. A mathematical study shows the existence of travelling waves. Then we provide a detailed numerical study of the system.

Finally, the last part, concerns the system of Patlak–Keller–Segel 1D after blow–up. The mathematical study provide a description of the system after blow–up, based on the notion of default meausure. Then we propose a numerical scheme, adopting the optimal transport viewpoint and allowing to simulate the system after blow–up.

Keywords: biology, metastasis, McKendrick–Von Foerster, respiration, Navier–Stokes, dissipative condtitions, screening effects, MAP Kinase, wave propagation, Keller–Segel, default measure, optimal transport.