UNIVERSITÉ DE RENNES 1

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1
COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Mathématiques et leurs Interactions*

Par
Bastien HAMLAT

Modélisation mathématique de réactions cinétiques multiphasiques en géochimie

Thèse présentée et soutenue à Rueil-Malmaison, le 22 septembre 2020
Unité de recherche : IRMAR, INRIA Rennes, IFP Energies nouvelles

Rapporteurs avant soutenance :

Etienne AHUSBORDE	Chargé de recherche, Université de Pau
Juliette LEBLOND	Directrice de recherche, INRIA Sophia Antipolis

Composition du Jury :

Présidente :	Danielle HILHORST	Directrice de recherche émérite, CNRS Orsay
Rapporteurs :	Etienne AHUSBORDE	Chargé de recherche, Université de Pau
	Juliette LEBLOND	Directrice de recherche, INRIA Sophia Antipolis
Examinateurs :	Mounir HADDOU	Professeur, INSA Rennes
	Michel KERN	Chargé de recherche, INRIA Paris
Directrice de thèse :	Jocelyne ERHEL	Directrice de recherche émérite, INRIA Rennes et IRMAR
Encadrants de thèse :	Anthony MICHEL	Ingénieur de recherche, IFP Energies nouvelles
	Thibault FANEY	Ingénieur de recherche, IFP Energies nouvelles

Remerciements

Une page se ferme en écrivant les dernières lignes de ce manuscrit... Je souhaite à tous les futurs étudiants doctorants d'arriver à s'épanouir pendant et après la thèse. Je garde en souvenir une tendresse pour chacune de mes rencontres pendant ces années de recherche et de rédaction.

Jocelyne ERHEL, Anthony MICHEL et Thibault FANEY, je tiens à vous exprimer mes plus sincères remerciements pour votre implication durant la thèse. Je vous remercie de m'avoir accordé votre confiance et beaucoup de liberté scientifique. Votre expérience et vos conseils m'ont été précieux. Vous avez tous fait part d'une parfaite bienveillance à mon égard et je vous en suis reconnaissant. Jocelyne, je te remercie d'avoir accepté de diriger ma thèse. Ce fut un réel plaisir d'apprendre à te connaître et de travailler avec toi. Mes visites à Rennes ont toujours été des moments agréables et très fructueux scientifiquement. Tes retours et tes apports m'ont permis de progresser au cours de ce long travail. Je garderai un très bon souvenir de nos échanges. Anthony, je te remercie de m'avoir recruté et pour avoir eu confiance en moi dès notre première rencontre. Notre travail collaboratif parfois quotidien a été moteur dans cette thèse. Tu as toujours été disponible lorsque j'avais besoin d'aide. Thibault, ton soutien et tes conseils scientifiques m'ont été indispensables durant la thèse. Je te remercie pour ta bonne humeur communicative, surtout durant les périodes difficiles.

Pour avoir accepté de faire partie du jury de la thèse je souhaiterais aussi remercier sincèrement la présidente du jury Danielle HILHORST, les rapporteurs Étienne AHUSBORDE et Juliette LEBLOND pour leurs retours qui ont permis l'amélioration du manuscrit et les examinateurs Mounir HADDOU et Michel KERN pour l'intérêt qu'ils ont accordé à mes travaux de thèse.

Je remercie Jocelyne de m'avoir présenté Tangi MIGOT et de nous avoir permis de travailler ensemble. J'ai eu un réel plaisir à travailler avec toi, tu es toujours enthousiaste et la difficulté ne te fait pas peur bien au contraire. Merci d'avoir partagé avec moi tes compétences scientifiques. Bien évidement, je te remercie d'avoir travaillé à mes côtés et je te souhaite le meilleur pour ta carrière scientifique. Et puisque tout est permis : que Rennes soit champion de France.

Durant ma thèse, j'ai eu le plaisir et la chance de faire mes entretiens annuels d'avancement du projet de thèse, organisés par l'école doctorale MathSTIC que je remercie très sincèrement, avec Michel KERN et Jean-Raynald DE DREUZY. Cela était bien plus qu'un simple entretien où l'on présente ses travaux. Nos discussions m'ont permis d'explorer des axes différents dans mes travaux et parfois de me rassurer lorsque les inquiétudes me tenaillaient.

Ma thèse est une collaboration entre IFP Energies nouvelles et INRIA. Je souhaite remercier IFP Energies nouvelles de m'avoir recruté en thèse, de m'avoir octroyé tout le confort matériel et financier dont a besoin un jeune chercheur pour effectuer son travail. Je remercie également INRIA de Rennes d'avoir financé plusieurs de mes déplacements et

de m'avoir toujours parfaitement bien accueilli lors de mes venues. Je souhaite remercier l'école doctorale MathSTIC. Merci à l'université Paris 2 Panthéon-Assas d'avoir financé la dernière année de ma thèse en m'octroyant un contrat ATER.

Je souhaite remercier la région Bretagne pour le financement de mon voyage à Milan dans le cadre de mes travaux de recherche. Je remercie le MOX de Politecnico di Milano pour m'avoir accueilli dans leur laboratoire. Merci à Luca FORMAGGIA et Luca BONAVENTURA de m'avoir permis de travailler dans leur équipe. Je remercie plus particulièrement Anna SCOTTI qui m'a supervisé pendant mon séjour (mathématiques, activités, problème de dents,...). J'ai été chanceux de la rencontrer, un grand merci. Je souhaite remercier Abramo AGOSTI pour les discussions scientifiques et l'aide qu'il a pu m'apporter. Une pensée pour tous les stagiaires, doctorants et post-doctorants : Federica, Ludovica, Francesco, Fannie, Laura, les deux Daniele, Alberto, Mattia, Matteo, Abele. Nous avons passé tant de bons moments ensemble (aperitivo + football) et vous m'avez même appris à faire le Tiramisu ! Enfin, je remercie Florent CHAVE, mon compère Montpelliérain, sans toi l'aventure Milanaise n'aurait pas été la même. Je te souhaite de réussir ce que tu entreprends.

Je tiens à remercier les membres de OSUR pour la summer school de Cargèse et plus particulièrement les doctorants que j'ai eu le plaisir de revoir plusieurs fois lors de mes séjours rennais.

Je souhaite remercier Terence BAYEN pour les discussions scientifiques que l'on a eu durant le congrès SMAI.

J'ai été parfaitement accueilli dans l'équipe Fluminance de INRIA Rennes. Merci à Etienne MEMIN le chef de projet. Je tiens particulièrement à remercier Huguette BECHU, avec qui j'ai beaucoup échangé, pour sa disponibilité, son efficacité et d'avoir fait de son mieux pour me permettre de me déplacer dans les meilleures conditions. Merci aux doctorants, post-doctorants et permanents de l'équipe pour les moments passés ensembles.

Je remercie tous les membres de R115 et R114. Une pensée particulière pour Zakia BENJELLOUN TOUIMI qui a accordé une réelle attention à ma thèse et qui a été pour moi un soutien indispensable, simplement merci. Je tiens à remercier Carole GUENAUD et Sylvie HOGUET qui m'ont toujours aidé et qui font un travail extraordinaire. Elles ont toujours été bienveillantes et agréables avec moi. Sylvie, je garde un souvenir particulier de nos discussions, j'ai appris à connaître une personne formidable. Je remercie Huy TRAN QUANG pour ses conseils scientifiques et son aide, les doctorants de R115/114 ont de la chance de pouvoir compter sur toi. J'ai une pensée pour chaque membre permanent de IFP Energies nouvelles que j'ai eu le plaisir de rencontrer : Abdallah, Ani, Anthony, Aziz, Benjamin, Bruno, Catherine, Chakib, Christophe, Delphine, Elsie, Françoise, Frédéric, Guillaume, Henry, Ibtihel, Isabelle, Jef, Jean-Marc, Jean-Yves, Jean-Louis, Julien, Huy, Léo, Long, Laurent, Mathieu, Miguel, Nina, Olivier, Paul, Pierre, Raphaël, Soleiman, Stéphane, Sylvain, les trois Sylvie, Thibault, les deux Thomas, Xavier, Zakia. C'était un plaisir de partager des instants avec chacun d'entre vous pendant les repas ou durant les

pauses de convivialité.
J'ai eu le plaisir d'animer le club Belote de IFP Energies nouvelles et pour cela je remercie mes collègues et le CE pour nous avoir permis de nous réunir tous les vendredi midi pour jouer à la belote. J'y ai rencontré des gens formidables et très sincèrement vous m'avez permis de tenir dans les moments difficiles. Vous êtes et resterez dans mon cœur, sachez que même si nos routes se séparent vous pourrez toujours compter sur moi. En tant que premier président du club, je vous demande de ne jamais arrêter ! Jean-Yves comment oublier nos discussions sur la musique, notre goût pour Motörhead nous a tout de suite rapprochés, tu m'as quand même fait redécouvrir Scorpions. Cela m'a touché quand tu m'as invité à joueur aux cartes avec ta famille dans ma région natale. Je suis fier de te compter parmi mes proches. Christophe, j'ai adoré parler de tout avec toi (musique, démineur, maths,...). Nous avons partagé ensemble notre passion commune des problèmes mathématiques et cela était très plaisant, tu es une personne formidable et j'ai aimé chaque instant passé à tes cotés. Sylvie P., cette page ne suffirait pas pour te témoigner ma gratitude. Tu as été un modèle qui m'a permis d'avancer. Tu es une de mes plus belles rencontres durant la thèse. Delphine, la « maman » des thésards. Tu es un bonheur quotidien. Ta bonne humeur et ta joie de vivre m'ont fait passé tellement de moments agréables durant cette thèse. Bruno, l'indispensable à la bonne ambiance. J'ai adoré jouer à MTG avec toi et discuter de tout. Tu es une super rencontre. Thomas, j'adore ton humour. Franchement tu es au top, ne change pas ! Xavier et Raphaël, le tandem à abattre à la coinche. Xavier, j'ai adoré discuter avec toi et rire de tout et de rien, tu es un gars super sympa. Raph, c'était vraiment un plaisir de partager pleins de moments avec toi, tu m'as beaucoup fait rire. Fred le Nîmois, nos racines communes gardoises nous ont rapprochés, j'ai passé de bons moments dans votre bureau (bureau de Delphine et Fred). Benoît, tu étais le seul à ne pas faire partie de notre département. Pourtant, tu nous as régalés avec tes règles bizarroïdes qui rendent un jeu que nous aimons tous différent mais encore plus fun ; tu n'étais pas non plus le dernier pour mettre l'ambiance. Long, ta bonne humeur me manquera. Sylvie W., je te remercie pour ton aide, j'ai eu beaucoup de chance de te compter parmi mes soutiens.

J'ai une pensé pour tous les stagiaires qui ont croisé ma route à IFPEN : Irene, Nawel, Anis, Lynda , Marianne, Imane, Hassan, Olivier, Jeremy,...
Une pensée pour Gouranga qui était en Post-doc à IFPEN. J'ai partagé beaucoup de moments avec toi et j'ai adoré nos discussions sur nos différences culturelles. Comment oublier mes collègues doctorants... Je tiens d'abord à tous vous dire que vous êtes de magnifiques rencontres. Mon année de recrutement a vu 3 doctorants être accueillis dans notre département : Julien, Karine et moi-même. Mes deux compères avec qui nous avons traversé tant d'épreuves. Karine, ce fut un plaisir de discuter avec toi. Nous avons passé des bons moments ensemble (sans parler de la pétanque !). Je te souhaite le meilleur pour la suite, tu le mérites. Julien, une sincère amitié est née entre nous. Tu fais partie des personnes importantes dans ma vie aujourd'hui. On a partagé un bureau pendant de longues années... mais j'espère qu'on continuera à partager une amitié tout le reste de notre existence. J'espère que tu t'épanouiras dans tes futures réalisations. Je tiens à remercier les doctorants qui m'ont accueilli. En particulier, Riad avec qui j'ai partagé tellement de bons moments « à la pétanque » entre autres. Zakariae, nous avons partagé le même bureau durant toute la durée de ma thèse. J'ai trouvé en toi un ami et tu fais par-

tie, aujourd'hui, des personnes importantes de ma vie. Tu es une personne exceptionnelle et je m'estime chanceux de t'avoir rencontré. Thói, j'ai eu le plaisir de partager le bureau avec toi. Nos échanges français/anglais et culturels étaient très enrichissants. Tu étais le « smile » de notre bureau. Tu es une très belle rencontre et je te souhaite le meilleur. Sabrina, tu as été pour moi la meilleure arrivée dans notre groupe de doctorants. Tu t'es parfaitement intégrée et tu as permis d'accroître la bonne ambiance collective. Toujours souriante, prête à aider, active pour organiser des activités, ... Tu es une personne sur qui on peut compter. Bien que je ne sois plus à IFP, je serai toujours là si tu as besoin. Guissel, tu as été une super rencontre. Ta joie de vivre et ton humour nous ont permis de passer des moments inoubliables. Alexis, pourtant un peu plus discret que les autres boute-en-trains, nous avons cependant beaucoup discuté et échangé sur des sujets divers et variés. Je te souhaite de réaliser tes futurs projets. Joëlle, j'ai aimé discuter avec toi. Tu es une super rencontre et je te souhaite sincèrement de réussir ce que tu entreprends. Son, nous n'avons que peu échangé mais c'était toujours un plaisir de parler avec toi. Ruben, chacune de tes venues était un plaisir. Narges, nous avons partagé une superbe expérience en Corse ensemble. Ce fut un plaisir de te rencontrer. Pour tous les doctorants qui n'ont pas encore soutenu je vous souhaite que ce jour arrive le plus rapidement possible et je suis sûr que vous ferez de bons docteurs.

Bien évidemment cette aventure n'aurait pas été la même sans mes proches. Une première pensée va à mes parents et mes sœurs. Mathilde, je suis extrêmement fier de toi. Tu es la sœur dont rêvent tous les frères. Marie, tu es encore jeune mais je te souhaite de réussir et surtout de t'épanouir. Madie, j'ai une pensée tendre pour toi. J'ai passé un agréable moment en ta compagnie lors de ma venue sur Lyon. Papa et Christelle, merci pour vos conseils et votre soutien. Maman et Yannick, merci pour votre entière dévotion et vos conseils. Maman merci pour l'amour que tu m'as apporté tout au long de ma vie. Je tiens à remercier mes grand-mères et les autres membres de ma famille. Une pensée particulière pour ma tante Annie qui est venue voir ma soutenance (cela m'a énormément touché). Je souhaite remercier Nicolas de m'avoir accueilli chez lui et de me considérer comme un membre de sa famille. Je suis triste de ne plus venir sur Rennes... Merci à mes amis proches qui m'ont soutenu durant cette épreuve : Alicia, Maïlys, Thibault, Julien, Olivier. Vous savez tous à quel point vous êtes importants dans ma vie.

Mes dernières pensées vont à la femme qui m'a épaulé, supporté et aimé. Amina nous y sommes arrivés... Je ne parle pas que de la thèse. Je suis heureux d'être celui que tu as choisi, tu combles de bonheur chacun de mes jours.

Résumé

Le sujet de thèse s'inscrit dans la continuité des travaux de thèse antérieurs réalisés à IFP Energies nouvelles ([42] et [68]) et à INRIA ([25]). Tous ces travaux tentent d'apporter des réponses en terme de modélisation mathématique et de méthode numérique pour la simulation du transport réactif en milieu poreux.

En transport réactif les apparitions et les disparitions de phases et d'espèces engendrent un problème de modélisation des cinétiques chimiques. Des travaux sur les cinétiques chimiques ont déjà été réalisés pour les systèmes chimiques comprenant une unique phase minérale pure par réaction. Notre objectif est d'élargir la gamme des systèmes chimiques que l'on peut modéliser et de proposer des méthodes numériques de résolution.

Nous considérons un système chimique avec une phase aqueuse et des phases pures, où les réactions chimiques sont cinétiques. Un modèle avec des vitesses de réaction continues convient en phase aqueuse mais ne convient pas dès qu'il y a une phase pure, car les quantités peuvent devenir négatives. Une difficulté est de maintenir les quantités de matière positives lorsqu'une espèce disparaît.

Dans le premier chapitre, un modèle de type EDOs discontinues pour la cinétique chimique avec apparitions et disparitions d'espèces pour un nombre quelconque de minéraux est proposé. Ce travail a été réalisé en collaboration avec Jocelyne Erhel, Thibault Faney et Anthony Michel. Une version régularisée du modèle permet de prouver la positivité et l'existence de solutions. De plus, une analyse explicite est effectuée dans le cas contenant une espèce réactive intermédiaire. Enfin, une méthode numérique basée sur la formulation régularisée a été validée sur des exemples de systèmes réactifs.

Dans le deuxième chapitre, une reformulation du modèle de cinétique chimique proposé dans le premier chapitre à l'aide de la théorie de Filippov est fournie. Une preuve de l'existence et de la positivité des solutions de la formulation inclusion différentielle ainsi obtenue est réalisé. De plus, dans le cas des surfaces de discontinuité de codimension 1, une étude des configurations des champs fournit un résultat d'unicité et de caractérisation des trajectoires.

Dans le troisième chapitre, un modèle de cinétique chimique de type systèmes dynamiques projetées est proposé. Ce travail est effectué en collaboration avec Jocelyne Erhel et Tangi Migot. Une analyse de l'existence des solutions de ce modèle, des liens avec d'autres types de formulations et une méthode de résolution numérique adaptée sont présentés. Enfin, une illustration des résultats numériques obtenus est réalisée pour des systèmes de cinétique chimique.

Mots-clés : cinétiques chimiques, EDOs discontinues, inclusions différentielles, systèmes dynamiques projetées, théorie de Filippov, simulations numériques, analyse de positivité et d'existence.

Abstract

The thesis is a continuation of the previous thesis work carried out at IFP Energies nouvelles ([42] et [68]) and at INRIA ([25]). All of these works attempt to provide answers to efficient modeling and numerical methods for resolving reactive transport in porous media.

In reactive transport, the appearance and disappearance of phases and species give rise to a problem in modeling chemical kinetics. Works on chemical kinetics has already been carried out for chemical systems comprising a single pure mineral phase per reaction. Our aim is to extend models and numerical methods to general chemical systems.

We consider a chemical system with an aqueous phase and pure phases, where the chemical reactions are kinetic. A model with continuous reaction rates is suitable in the aqueous phase but no longer appropriate as soon as there is a pure phase, because the quantities can become negative. One difficulty is to maintain the quantities of matter positive when a species disappears.

In the first chapter, a discontinuous ODEs model for chemical kinetics with appearance and disappearance of species for any number of minerals is proposed. This work is made in collaboration with Jocelyne Erhel, Thibault Faney and Anthony Michel. A regularized version of the model allows to prove the positivity and the existence of solutions. In addition, an explicit analysis is performed in a case with an intermediate reactive species. Finally, a numerical method based on the regularized formulation has been validated on examples of reactive systems.

In the second chapter, a reformulation of the chemical kinetics model proposed in the first chapter using Filippov's theory is provided. A proof of the existence and the positivity of the solutions of the differential inclusion formulation thus obtained is produced. In addition, in the case of the discontinuity surfaces of codimension 1, a study of the configurations of the fields provides a result of uniqueness and characterization of the trajectories.

In the third chapter, a model of chemical kinetics of the projected dynamical system is proposed. This work is carried out in collaboration with Jocelyne Erhel and Tangi Migot. An analysis of the existence of solutions of this model, a link with other types of formulations and an adapted numerical resolution method are provided. An illustration of the numerical results obtained is made for chemical kinetic systems.

Keywords : chemical kinetics, discontinuous ODEs, differential inclusions, projected dynamical systems, Filippov's theory, numerical simulation, analysis of positivity and existence.

Table des matières

Glossaire **1**

Introduction **5**
0.1 Modèle de transport réactif multiphasique 5
0.1.1 Description du système physique 5
0.1.2 Inconnues du système 6
0.1.3 Lois de fermeture 6
0.1.4 Physique de l'écoulement en milieu poreux 7
0.1.5 Interactions chimiques 8
0.1.6 Système d'équations 10
0.2 Etat de l'art 11
0.3 Descriptions des travaux de thèse 13
0.3.1 Chapitre 1 : Modèles de cinétique chimique avec des phases pures 14
0.3.2 Chapitre 2 : Analyse d'un modèle de type EDOs discontinues pour la cinétique chimique multiphasique 17
0.3.3 Chapitre 3 : Modèle avec dynamique projetée pour les réactions de dissolution et de précipitation des minéraux en géochimie 18

1 Kinetic models with pure phases **21**
1.1 Introduction 21
1.2 Mathematical model of kinetic reactions 22
1.2.1 Examples 22
1.2.2 Quantities of species and reaction rates 23
1.2.3 Conservation laws 23
1.2.4 Maximal reaction rates 24
1.2.5 Continuous model 25
1.3 State of the art: one mineral in one reaction 26
1.3.1 Discontinuous model 26
1.3.2 Set-valued model 27
1.3.3 Complementarity model 27
1.3.4 Reference example 28
1.3.4.1 Solution of the discontinuous model 28
1.3.4.2 Precipitation 29
1.3.4.3 Partial dissolution 29
1.3.4.4 Full dissolution 29
1.4 A differential inclusion model for general geochemical systems 30

1.4.1 Regularized model . . . 31
1.4.2 Convergence of the regularized solutions . . . 33
1.5 Test case with three minerals . . . 34
1.5.1 Dissolution of M_1 and M_3 . . . 35
1.5.2 Dissolution of M_2 . . . 36
1.5.3 Dissolution of M_1 then M_2 . . . 37
1.5.4 Dissolution of M_2 then M_1 . . . 39
1.5.5 Summary of test case $M_1 - M_2 - M_3$ . . . 40
1.6 Numerical experiments . . . 41
1.6.1 Numerical context . . . 41
1.6.1.1 ODE solver . . . 41
1.6.1.2 Physical model . . . 42
1.6.1.3 Discrete error norms . . . 42
1.6.2 Reference example $A - M$, total dissolution . . . 42
1.6.3 Test case $M_1 - M_2 - M_3$ . . . 45
1.6.4 Test case $A_1 - M_2 - A_3$ . . . 48
1.6.5 Test case with competition . . . 50
1.6.6 Test case with blending (non-uniqueness) . . . 52

2 Analysis of a discontinuous ODEs model for multiphase chemical kinetics 55
2.1 Introduction . . . 55
2.2 Differential inclusion model for chemical kinetics . . . 56
2.2.1 Domains and surfaces of discontinuity . . . 56
2.2.2 Extension of reaction rates . . . 57
2.2.3 Differential inclusion kinetic model . . . 58
2.2.4 Essential positivity property . . . 60
2.2.5 Existence of solutions . . . 61
2.3 Uniqueness of the solution when only one mineral disappears . . . 63
2.3.1 Attractive field . . . 65
2.3.2 Transverse field . . . 68
2.3.3 Tangent field . . . 69

3 A projected dynamical system approach to mineral precipitation-dissolution reactions in geochemistry 73
3.1 A Projected Dynamical System Model . . . 75
3.1.1 Kinetic system . . . 75
3.1.2 Model . . . 75
3.2 Link with Other Models . . . 77
3.2.1 Link with a complementarity system . . . 77
3.2.2 Link with a discontinuous model . . . 78
3.2.3 Link with a classical projected dynamical system . . . 79
3.3 Existence of a Solution . . . 80
3.3.1 Explicit scheme and interpolation . . . 80
3.3.2 Bounded constraints . . . 81
3.3.3 Convergence of the discrete function . . . 82

3.3.4 Existence of a solution 84
3.4 An Alternating Projection Algorithm 85
3.5 Numerical Experiments 85
3.5.1 Example W_1W_2 86
3.5.2 Example WM 87
3.5.3 Example $M_1M_2M_3$ 87
3.5.4 Example WM_1M_2 88
3.6 Concluding Discussion 89

Conclusion et perspectives **91**

Bibliography **95**

Glossaire

Notions et notations des chapitres 1 et 2

— Espèces chimiques :
aqueuses A_i, $i = 1, \cdots, N_a$
minérales M_i, $i = 1 + N_a, \cdots, N_a + N_p$

— Réactions chimiques : $\mathcal{R}_j$, $j = 1, \cdots, N_r$

— Stœchiométrie : $S = (s_{i,j})_{\substack{1 \leqslant i \leqslant N_a + N_p \\ 1 \leqslant j \leqslant N_r}}$ et $s_{i,j} \in \mathbb{Z}$

— Phase pure est une phase contenant une unique espèce (dans nos travaux il s'agit exclusivement des minéraux)

— Multiphasique : ce dit d'un système comprenant plusieurs phases (dans nos exemples il y a toujours une phase aqueuse et une ou plusieurs phases pures minérales)

— Réactifs de $\mathcal{R}_j$ (indices) : $\mathcal{A}_j^- = \{i \in \{1, \cdots, N_a\},\ s_{i,j} < 0\}$ et $\mathcal{M}_j^- = \{i \in \{1 + N_a, \cdots, N_a + N_p\},\ s_{i,j} < 0\}$

— Produits de $\mathcal{R}_j$ (indices) : $\mathcal{A}_j^+ = \{i \in \{1, \cdots, N_a\},\ s_{i,j} > 0\}$ et $\mathcal{M}_j^+ = \{i \in \{1 + N_a, \cdots, N_a + N_p\},\ s_{i,j} > 0\}$

— Réaction $\mathcal{R}_j$ (voir (1.1)) :

$$\sum_{i \in \mathcal{A}_j^-} |s_{ij}| A_i + \sum_{k \in \mathcal{M}_j^-} |s_{kj}| M_k \rightleftarrows \sum_{i \in \mathcal{A}_j^+} |s_{ij}| A_i + \sum_{k \in \mathcal{M}_j^+} |s_{kj}| M_k.$$

— Quantités (concentrations) : $\boldsymbol{c} = (c_i)_{1 \leqslant i \leqslant N_a + N_p} = (\boldsymbol{c_a}, \boldsymbol{c_p})$ avec comme contrainte à respecter $\boldsymbol{c} \geqslant 0$
espèces aqueuses $\boldsymbol{c_a} = (c_i)_{1 \leqslant i \leqslant N_a}$
espèces minérales $\boldsymbol{c_p} = (c_i)_{1 + N_a \leqslant i \leqslant N_a + N_p}$

— Activités :
minéraux $a_i = 1$, $1 + N_a \leqslant i \leqslant N_a + N_p$
espèces aqueuses $a_i(\boldsymbol{c_a})$, $1 \leqslant i \leqslant N_a$ avec (Assumption 1.2.1)

$$\mathbf{c}_a \geqslant 0 \Rightarrow \mathbf{a}(\mathbf{c}_a) \geqslant 0$$

$$c_{a,i} = 0 \Rightarrow a_i(\mathbf{c}_a) = 0.$$

— Loi d'action de masse ou taux "maximal" de réaction : τ_j pour la réaction $\mathcal{R}_j$, $j = 1, \cdots, N_r$ (voir (1.3))

$$\tau_j(\mathbf{c}_a) = \alpha_j \prod_{i \in \mathcal{A}_j^-} a_i(\mathbf{c}_a)^{|s_{ij}|} - \beta_j \prod_{i \in \mathcal{A}_j^+} a_i(\mathbf{c}_a)^{|s_{ij}|}, j = 1, \ldots, N_r.$$

— Vitesse de réaction : r_j pour la réaction $\mathcal{R}_j$, $j = 1, \cdots, N_r$

— Système cinétique (voir (1.2))
conservation de la masse $\dfrac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} = S\mathbf{r}(\mathbf{c})$
contrainte de positivité des quantités $\mathbf{c} \geqslant 0$,
condition initiale $\mathbf{c}(0) = \mathbf{c}_0$

— Invariants : Q base de $\ker(S^T)$ ($S^T Q = Q^T S = 0$)
$\mathbf{c}$ solution (satisfait la conservation de la masse) implique $Q^T \mathbf{c} = Q^T \boldsymbol{c_0}$

— Modèle continue : avec $\boldsymbol{r} = \boldsymbol{\tau}$ (voir (1.4))

— Limiteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ (voir (1.14))

— Modèle limité (inclusion différentielle) (voir (1.15))

— Modèle régularisé (voir (1.19))

— Blending : la fonction de mélange des limiteurs pour le modèle régularisé (voir (1.18) et (1.27))

— Système réactif avec compétition : Si la consommation d'une espèce du système est due à plusieurs réactions du système

— Surfaces de discontinuité : $\Sigma = \bigcup_{i \in \{1, \cdots, N_p\}} \Sigma_i$
avec $\Sigma_i = \{\boldsymbol{c} = (\boldsymbol{c_a}, \boldsymbol{c_p}) \in \mathbb{R}^{N_a + N_p} \mid c_{p,i} = 0\}$ les hyperplans correspondant à la valeur de concentration nulle d'un minéral

— Domaines : G_k le plus grand ensemble ouvert caractérisé par une valeur unique du signe des concentrations des minéraux

— Vecteur représentant de G_k : $\boldsymbol{e^k}$

$$e_i^k = \begin{cases} 1, & \text{if } \forall \boldsymbol{c} \in G_k,\ c_{p,i} > 0, \\ 0, & \text{if } \forall \boldsymbol{c} \in G_k,\ c_{p,i} < 0. \end{cases}$$

— Ensemble des indices des domaines voisins à un point $\boldsymbol{c}$: $\mathcal{G}(\boldsymbol{c})$

— Modèle de type Filippov (inclusion différentielle) (voir (2.8))

— Codimension d'un point $\boldsymbol{c}$: le nombre de composente minérale égale à zéro

— Champs attractif : voir (2.15)

— Champs transverse : voir (2.17)

— Champs tangent : voir (2.18)

Notions et notations du chapitre 3

— Espèces chimiques : nombre d'espèces totales n
aqueuses W_i et minérales M_i

— Nombre de réactions : m

— Stœchiométrie : $S \in \mathbb{Z}^{n\times m}$

— Quantités (concentrations) : x

— Espèce aqueuses (ensemble d'indice) : $\mathcal{W}$

— Taux de réaction : τ (voir (3.1))

— Invariants : Q base de $\ker(S^T)$ $(S^TQ = Q^TS = 0)$

— Dérivée : $\dfrac{\mathrm{d}x}{\mathrm{d}t} = \dot{x}$

— Projection : $\mathrm{Proj}(C; v) := \{w \in C : \|v - w\|_2 = \inf_{z\in C}\|v - z\|_2\}$

— Cône tangent :(voir (3.2))

$$\mathcal{T}(C; x) = \{d : \exists d_\nu \to d \text{ and } \exists \tau_\nu \downarrow 0 \text{ with } x + \tau_\nu d_\nu \in C\ \forall \nu\}.$$

— Contraintes sur les concentrations : $C_0 = \{x \in \mathbb{R}^n_+ \ : \ Q^T(x - x_0) = 0\}$

— Contraintes sur la vitesse des concentrations : $\mathcal{T}(C_0; x) \cap D(x)$
avec $D(x) := \{\mathrm{diag}(S\tau(x))\lambda \text{ with } \lambda \in [0, 1]^n\}$
— Système avec dynamique projetée : voir (3.4)

Introduction

Cette thèse s'inscrit dans la continuité des travaux de thèse antérieurs réalisés à IFP Energies nouvelles ([42] et [68]) et à l'INRIA ([25]). Ces travaux tentent d'apporter une réponse aux difficultés liées à la modélisation et à la simulation numérique du transport réactif en milieu poreux. Ce domaine a un grand nombre d'applications industrielles : EOR-CO_2 (production assistée d'hydrocarbures par injection de CO_2), stockage du gaz (injection de gaz dans le sous-sol afin de le conserver et de l'utiliser pour répondre à une demande ultérieure) ou encore géothermie (circulation d'un fluide dans le sous-sol afin de récupérer de l'énergie sous forme de chaleur). Ces applications sont basées sur l'écoulement d'une ou plusieurs phases fluides (eau, gaz, huile) dans une roche poreuse composée de phases minérales pouvant engendrer des interactions fluide-roche. Il est ainsi essentiel de disposer de formulations mathématiques bien posées et de solveurs numériques efficaces permettant de modéliser la complexité physique des systèmes étudiés.

0.1 Modèle de transport réactif multiphasique

Dans cette thèse, nous proposons des modèles mathématiques permettant de modéliser la cinétique chimique des interactions fluide-roche. Nous présentons ici la structure d'un problème de transport réactif multiphasique. L'objectif est de décrire les inconnues et les paramètres du système, ainsi que les équations à résoudre.

0.1.1 Description du système physique

Dans l'étude du transport réactif multiphasique en milieu poreux, il faut distinguer les lois et paramètres physiques issus de la physique de l'écoulement et ceux issus des interactions chimiques. Le système compositionnel d'un milieu poreux, noté Σ, permet de décrire la composition chimique du milieu poreux. Il est composé de deux sous-systèmes caractérisés par les états de matière : le sous-système fluide $\mathcal{F} \in \Sigma$, comprenant les liquides et les gaz, et le sous-système solide $\mathcal{S} \in \Sigma$, comprenant les solides. Chaque sous-système est lui-même constitué de phases notées α qui sont des ensembles de constituants chimiques pour lesquels les propriétés physiques sont homogènes. Les phases du sous-système solide (i.e. $\alpha \in \mathcal{S}$) considérées dans ce document sont pures, c'est à dire qu'elles ne contiennent qu'un unique constituant chimique. En d'autre termes, chaque minéral constitue une phase du sous-système solide. Chaque phase est composée de plusieurs constituants chimiques appelés espèces chimiques. On note $\mathcal{E}$ l'ensemble des espèces chimiques du système compositionnel Σ et $\mathcal{E}_\alpha$ l'ensemble des espèces chimiques contenues dans la phase α. Lorsqu'on utilise des espèces abstraites $E \in \mathcal{E}$, au sens où l'on ne

connaît pas leur constitution atomique, on utilise la convention suivante : W (ou A) complété d'un indice dénote une espèce en phase aqueuse (i.e. phase dans laquelle l'eau est le solvant), M complété d'un indice dénote un minéral pur et G complété d'un indice dénote une espèce gazeuse. Par conséquent, une fois ordonnée, une espèce dans une phase est entièrement caractérisée par son indice dans la phase, cela permet d'utiliser la notation $i \in \alpha$ pour référencer la i-ème espèce dans la phase α.

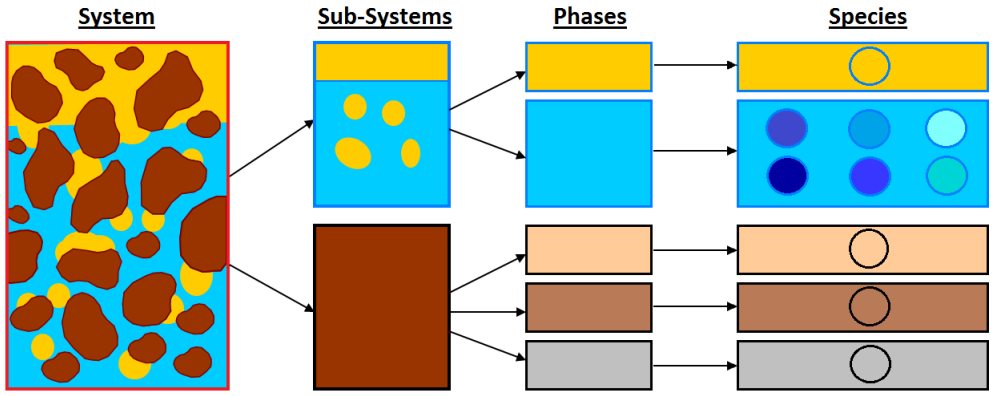


Figure 1 – Description du système compositionnel

0.1.2 Inconnues du système

Les inconnues du problème de transport réactif, détaillées plus loin dans cette section, sont un ensemble de variables physiques et chimiques caractérisant, à chaque instant et dans tout le domaine, la composition et les propriétés du système Σ. Il y a sept types d'inconnues différents. Le premier est la pression notée P (unité Pa). Le deuxième est la température notée T (unité K). Le troisième est la fraction volumique du fluide notée $\phi_{\mathcal{F}}$ (sans unité), elle représente la proportion de fluide dans le système compositionnel (i.e. la fraction entre le volume du fluide et le volume total du système). Le quatrième est la fraction volumique du solide notée $\phi_{\mathcal{S}}$ (sans unité), elle représente la proportion de solide dans le système compositionnel (i.e. la fraction entre le volume du solide et le volume total du système). Le cinquième est la saturation d'une phase fluide α notée S_α (sans unité). Le sixième est la fraction volumique d'une phase solide α notée F_α (sans unité). Le septième est la fraction molaire des espèces chimiques, on note x_i^α (sans unité) la fraction molaire de la i-ème espèce dans la phase α. Au total, le nombre d'inconnues est égal à la somme du cardinal de l'ensemble des espèces et du cardinal de l'ensemble des phases, à laquelle il faut ajouter les quatre inconnues restantes (pression, température, fraction volumique du fluide et du solide).

0.1.3 Lois de fermeture

Avant de préciser les équations qui modélisent les phénomènes physiques et chimiques dans un milieu poreux, on introduit les équations de fermeture. Ces équations sont structurelles, elles permettent de garantir la cohérence du système d'inconnues.

Equation de fermeture du système :

$$\phi_{\mathcal{S}} + \phi_{\mathcal{F}} = 1, \tag{1}$$

l'équation de fermeture du système (1) permet de maintenir constante la somme de la proportion de fluide et de solide dans le système.

Equation de fermeture du fluide :

$$\sum_{\alpha \in \mathcal{F}} S_\alpha = 1, \tag{2}$$

Equation de fermeture du solide :

$$\sum_{\alpha \in \mathcal{S}} F_\alpha = 1, \tag{3}$$

les équations de fermeture du fluide (2) et du solide (3) permettent de maintenir la somme des proportions des phases qui les composent à 1.

Equations de fermeture des phases :

$$\sum_{i \in \alpha} x_i^\alpha = 1, \ \forall \alpha \in \mathcal{S} \cup \mathcal{F}, \tag{4}$$

les équations de fermeture des phases (4) permettent de maintenir la somme des fractions molaires de chaque phase à 1. Le nombre d'équations de fermeture correspond au cardinal de l'ensemble des phases auquel on additionne trois. Il reste donc le cardinal de l'ensemble des espèces plus une équation à ajouter pour égaliser le nombre d'inconnues.

0.1.4 Physique de l'écoulement en milieu poreux

Les problèmes de transport réactif en milieu poreux font intervenir plusieurs lois physiques qui modélisent l'écoulement du fluide. Le système d'équations résultant est un système d'équations différentielles aux dérivées partielles (EDP). Il s'agit plus précisément d'une équation d'évolution avec un terme d'advection et de diffusion. Les paramètres suivants sont nécessaires pour définir la vitesse de Darcy [24] dans la phase α notée v_α ($\mathrm{m} \cdot \mathrm{s}^{-1}$) : le tenseur de perméabilité absolu K (m^2), la perméabilité relative associée à une phase kr_α (sans unité), la viscosité d'une phase μ_α ($\mathrm{Pa} \cdot \mathrm{s}$), l'accélération de la gravité g ($\mathrm{m} \cdot \mathrm{s}^{-2}$), la masse volumique d'une phase ρ_α ($\mathrm{kg} \cdot \mathrm{m}^{-3}$), la masse molaire d'une phase M_α ($\mathrm{kg} \cdot \mathrm{mol}^{-1}$) et la pression capillaire d'une phase Pc_α (Pa).

On peut définir la loi de Darcy comme suit, $\forall \alpha \in \mathcal{F}$,

$$v_\alpha = -K \frac{kr_\alpha}{\mu_\alpha} \big(\nabla (P + P_{c,\alpha}) - \rho_\alpha g \big). \tag{5}$$

Un exemple pour les lois que suivent la perméabilité relative et la pression capillaire sont les lois d'hydrostatique de Brooks-Corey et sont définies dans le cas biphasique (phase aqueuse w et phase gazeuse g) comme suit,

$$\overline{S}_w = \frac{S_w - S_{w,r}}{1 - S_{w,r} - S_{g,r}}, \tag{6}$$

où $S_{\alpha,r}$ est la saturation résiduelle de la phase $\alpha \in \{w, g\}$.

$$P_{c,g} = P_{e,g} \, (\overline{S}_w)^{-\frac{1}{\lambda}}, \tag{7}$$

$$kr_w = \quad \overline{S}_w^{\frac{2+3\lambda}{\lambda}}, \tag{8}$$

$$kr_g = \quad (1 - \overline{S}_w)^2 \; \overline{S}_w^{\frac{2+\lambda}{\lambda}}, \tag{9}$$

où $P_{e,g}$ est une constante de pression et λ est égal à 2 dans la plupart des cas.

La loi de Darcy permet de définir la vitesse d'advection d'un problème de transport réactif en milieu poreux. L'opérateur de transport L, dépendant de la concentration des espèces chimiques, est défini pour chaque espèce dans une phase. Il est noté L_i^α pour la i-ème espèce de la phase α. On note c_i^α ($\mathrm{mol} \cdot \mathrm{m}^{-3}$) la concentration de la i-ème espèce dans la phase α, cette grandeur est proportionnelle à la fraction molaire de l'espèce considérée,

$$c_i^\alpha = x_i^\alpha \frac{\rho_\alpha}{M_\alpha}, \tag{10}$$

et on note D_i^α ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$) le coefficient de diffusion de la i-ème espèce dans la phase α.

Opérateur de transport advectif et diffusif :

$$L_i^\alpha(c_i^\alpha) \;=\; \mathrm{div}\big(c_i^\alpha v_\alpha - D_i^\alpha \nabla c_i^\alpha\big), \quad \alpha \in \mathcal{F}, \tag{11}$$

$$L_i^\alpha(c_i^\alpha) \;=\; 0, \qquad\qquad\quad \alpha \in \mathcal{S}. \tag{12}$$

Un autre phénomène physique à prendre en compte est l'apport de matière dû aux puits (producteurs et injecteurs) notés $p \in \mathcal{P}$, ce terme source qui vient s'additionner au transport est noté $\mathcal{Q}$. Le terme source $\mathcal{Q}$ est défini par espèce, on le note $\mathcal{Q}_i^\alpha$, il est dépendant de la pression et de la concentration.

Termes sources de type puits :

$$\mathcal{Q}_i^\alpha = \sum_{p \in \mathcal{P}} \big(c_{i,p}^\alpha T_{p,\alpha}(P + Pc_\alpha - P_{p,\alpha}) + d_{p,i}(c_i^\alpha - c_{p,i}^\alpha) \big), \tag{13}$$

tel que $c_{i,p}^\alpha$ ($\mathrm{mol} \cdot \mathrm{m}^{-3}$) est la concentration de la i-ème espèce dans le p-ème puits, $T_{p,\alpha}$ est la transmissivité, $P_{p,\alpha}$ (Pa) est la pression de sa phase α et $d_{p,i}$ est un coefficient de diffusivité.

0.1.5 Interactions chimiques

Les interactions chimiques considérées sont des réactions chimiques de deux natures différentes : les réactions d'équilibre et les réactions de cinétique. Les réactions d'équilibre modélisent des interactions chimiques instantanées par rapport à l'échelle de temps considérée du phénomène physique. Mathématiquement, les considérer revient à résoudre un problème algébrique issu d'une élimination des variables dites secondaires du système d'EDP initial. Pour plus de détails se reporter aux différents travaux concernant les tables de Morel [60]. Les cinétiques modélisent les transferts chimiques lents. Mathématiquement, la prise en compte des cinétiques se fait par l'ajout d'un terme source de vitesse de réaction au système de transport. Ces deux types d'interactions peuvent coexister et

mathématiquement il n'y a aucune contradiction à considérer les deux simultanément. En effet, il suffit d'inclure les seconds membres liés à la cinétique puis d'effectuer l'élimination des variables et l'ajout des équations algébriques pour les équilibres.

Le système réactif est constitué de l'ensemble des réactions chimiques du problème. Une réaction chimique modélise les transformations de la matière. Dans la suite, l'indice i est rattaché à l'espèce et l'indice j à la réaction. Une réaction chimique se définit par une formule mathématique associant les espèces chimiques aux coefficients de stœchiométrie notés $s_{i,j}$. Le coefficient stœchiométrique est une proportion d'une espèce consommée ou produite dans une réaction.

Les réactions d'équilibre sont symbolisées par le signe = tandis que les réactions cinétiques, considérées comme réversibles, par le signe $\rightleftharpoons$.

Exemple 0.1.1 (Système réactif)**.**

$$2W_1 + W_2 = M_1 \tag{14}$$
$$W_3 = G_1 \tag{15}$$
$$3W_4 + M_2 \rightleftharpoons M_3 \tag{16}$$

Ce système réactif est composé de 8 espèces dont 4 espèces aqueuses, 3 espèces minérales et 1 espèce gazeuse. Il est constitué de 3 réactions dont les deux premières sont des réactions d'équilibre et la dernière est une réaction de cinétique.

Pour définir le processus réactionnel les constantes d'équilibre, notées K_i, pour la i-ème espèce, doivent être fournies. La vitesse de chaque réaction cinétique j est notée τ_j. Les vitesses de réaction sont modélisées par la loi d'action de masse cinétique :

$$\tau_j = k_{r,j} \prod_{\substack{i\in\{1,\cdots,n\}\\ s_{i,j}<0}} [E_i]^{|s_{i,j}|} - k_{p,j} \prod_{\substack{i\in\{1,\cdots,n\}\\ s_{i,j}>0}} [E_i]^{|s_{i,j}|}. \tag{17}$$

où $k_{p,j}$, $k_{r,j}$ sont des constantes positives dépendantes des constantes d'équilibre des espèces intervenant dans la réaction et de la constante de vitesse de la réaction. $[E_i]$ sont les activités des espèces chimiques.

L'activité d'une espèce chimique est une mesure de l'écart des propriétés de l'espèce chimique entre un état standard et un mélange réel à la même température. Cette quantité est modélisée par une loi fonction de la fraction molaire,

$$[E_i] = \gamma_i x_i, \tag{18}$$

où γ_i est le coefficient d'activité de l'espèce chimique E_i. L'activité des espèces en phase pure est égale à 1.

Les réactions chimiques peuvent être représentées par la matrice de stœchiométrie S. Cette matrice contient les réactions chimiques. Pour cela, elle stocke en colonne les coefficients stœchiométriques de chaque espèce dans la réaction considérée. Ces coefficients sont signés par leur présence à droite ou à gauche de la réaction. Pour la suite on considérera la convention suivante, les espèces réactives (présentes à gauche de la réaction) auront des coefficients négatifs et les espèces produites (présentes à droite de la réaction) auront des coefficients positifs. La matrice de stœchiométrie est la concaténation de la matrice S_{kin}, matrice de stœchiométrie pour les réactions de cinétique, et de la matrice S_{eq}, matrice de stœchiométrie pour les réactions d'équilibre, $S = (S_{eq}|S_{kin})$.

Exemple 0.1.2 (Matrice de stœchiometrie)**.** *La matrice du système réactif de l'exemple 0.1.1 est définie comme suit, on choisit de ranger les espèces de la manière suivante (W_1, W_2, W_3, W_4, M_1, M_2, M_3, G_1),*

$$S_{eq} = \begin{pmatrix} -2 & 0 \\ -1 & 0 \\ 0 & -1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \; S_{kin} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -3 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \; and \; S = \begin{pmatrix} -2 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

La matrice de stœchiométrie nous permet de déduire les quantités de matière conservées. La matrice des quantités de matière conservées est notée Q et se déduit du noyau de la transposé de S : $vec(Q) = \ker(S^T)$. Respectivement, on définit Q_{kin} et Q_{eq} de la même manière.

Nous allons préciser la contribution des réactions d'équilibre, comme énoncé en début de section ces réactions contribuent par l'ajout d'un système algébrique à résoudre. Pour cela il suffit de définir les contraintes à respecter. Elles dépendent du potentiel chimique, $\boldsymbol{\mu}$ ($\mathrm{J} \cdot \mathrm{mol}^{-1}$).

$$\mu_i = \mu_i^0 + RT \ln([E_i]), \tag{19}$$

où μ_i^0 est le potentiel chimique de l'espèce E_i dans l'état standard à la même température.

Le système d'équations induit par les équilibres chimiques s'obtient en multipliant les équations de conservations par la matrice Q_{eq} et par l'ajout des contraintes linéaires $S_{eq}\boldsymbol{\mu} = 0$.

0.1.6 Système d'équations

Dans l'objectif de définir le système d'équations modélisant le transport réactif multiphasique, nous définissons, ci-dessous, le vecteur des quantités de matière $\boldsymbol{n}$ (mol),

$$n_i^\alpha = \phi_{\mathcal{F}} \, S_\alpha \, \frac{\rho_\alpha}{M_\alpha} \, x_i^\alpha, \quad \alpha \in \mathcal{F}, \tag{20}$$

$$n_i^\alpha = \phi_{\mathcal{S}} \, F_\alpha \, \frac{\rho_\alpha}{M_\alpha} \, x_i^\alpha, \quad \alpha \in \mathcal{S}. \tag{21}$$

En assemblant tous les concepts définis précédemment, nous définissons le système d'équations aux dérivées partielles et algébriques modélisant le transport réactif multi-

phasique comme suit :

$$\phi_{\mathcal{S}} + \phi_{\mathcal{F}} = 1, \tag{22}$$

$$\sum_{\alpha \in \mathcal{F}} S_\alpha = 1, \tag{23}$$

$$\sum_{\alpha \in \mathcal{S}} F_\alpha = 1, \tag{24}$$

$$\sum_{i \in \alpha} x_i^\alpha = 1, \quad \forall \alpha \in \mathcal{S} \cup \mathcal{F}, \tag{25}$$

$$Q_{eq} \left(\frac{\partial \boldsymbol{n}}{\partial t} + \boldsymbol{L}(\boldsymbol{c}) + \boldsymbol{\mathcal{Q}} + S_{kin} \boldsymbol{\tau} \right) = 0, \tag{26}$$

$$S_{eq}\, \boldsymbol{\mu} = 0. \tag{27}$$

0.2 Etat de l'art

Les formulations globalement implicites pour le transport réactif en milieu poreux sont apparues pour des problèmes de transport couplés avec des équilibres de phases. Ces formulations, voir [19] et [55], permettent de résoudre de manière implicite des problèmes de transport multiphasique avec des phases fluides mélanges et des équilibres de phases. Des formulations incluant les cinétiques ont été proposées dans [27], [26], [42], [45], [51] et [68]. Ces formulations globalement implicites permettent l'ajout d'un minéral, pour les équilibres et pour les cinétiques, en interaction chimique avec la phase mobile, fluide.

Les formulations de transport réactif multiphasique en milieu poreux sont des EDPs couplées avec des équations algébriques et des contraintes d'inégalités. Une liste non exhaustive des articles proposant un modèle de transport réactif, voir [11], [19], [40], [46], [49], [51], [65], [70], [71] et [72]. L'ajout du multiphasique aux transferts de matière entraîne des possibles disparitions et apparitions de phases. Le système couplé doit maintenir les quantités de matière positives à l'aide des contraintes d'inégalité. La gestion de ces inégalités pose de réelles difficultés dans les simulateurs. Des travaux sur la prise en compte de ces contraintes ainsi que sur l'efficacité et la robustesse des schémas numériques ont été réalisés dans le but de simuler numériquement le transport réactif [5], [14], [15], [17], [23], [27], [26], [36], [37], [42], [44], [45] [50], [52], [53], [55], [56], [57], [62], [66] et [68].

La chimie des milieu poreux est de nature complexe et variée. Pour une introduction à la géochimie on se réfère aux livres suivants: [10] et [54]. Les systèmes réactifs peuvent contenir des réactions de type cinétique et de type équilibre. Elles peuvent aussi contenir des espèces présentes dans différentes phases, par exemple des minéraux et des espèces aqueuses. Dans [6], l'auteur explique la diversité des systèmes réactifs et propose une représentation de ceux-ci. La vitesse des réactions cinétiques est modélisée par la loi d'action de masse [18] et [58]. Cette loi dans le cas de réaction purement aqueuse assure à elle seule la positivité des concentrations des espèces.

Dans cette thèse, on se focalise sur les difficultés engendrées par la prise en compte des réactions de cinétiques chimiques contenant des minéraux. Une des principales difficultés est de gérer l'apparition et la disparition des phases en maintenant la positivité des

concentrations des espèces. Dans le cas des équilibres chimiques, un travail récent [35] a notamment permis de caractériser les contextes de présence des minéraux et d'assurer les propriétés mathématiques du modèle proposé. En cinétique, des modèles de nature différente ont été développés pour garantir la positivité des concentrations lors d'un processus de dissolution ou de précipitation d'un minéral. A notre connaissance, ces formulations ne peuvent prendre en considération plus d'un minéral par réaction.

Dans [12], [44], [46], [51], [71] et [72], les auteurs proposent un modèle d'EDOs discontinues pour modéliser les réactions cinétiques. Ces modèles utilisent une vitesse de réaction modifiée de telle sorte que la réaction ne peut plus consommer le minéral lorsqu'il disparaît. La nouvelle vitesse de réaction est obtenue en multipliant le terme qui consomme le minéral par un indicateur de présence du minéral. Les auteurs démontrent l'existence, la positivité et l'unicité des solution par l'utilisation de résultats classiques d'analyse (Lemme de Grönwall et Théorème du point fixe).

Une autre alternative est fournie dans [44] et [46] : les auteurs proposent un modèle de type complémentarité pour gérer les apparitions et disparitions des espèces. Une nouveauté qu'apporte cette formulation est la gestion des contextes de présence des espèces directement dans le système d'équations. Cela permet l'utilisation de méthodes numériques plus efficaces [48] et [63] que celles basées sur la gestion de chaque contexte de présence.

Dans [2] et [44], les auteurs proposent une formulation de type inclusion différentielle pour les cinétiques chimiques. Dans [44], les auteurs démontrent l'équivalence des trois types de formulation (EDOs discontinue, complémentarité et inclusion différentielle) au sens faible. Dans [2], les auteurs utilisent des résultats classiques de la théorie des inclusions différentielles pour prouver l'existence et l'unicité de la solution de son modèle, ils utilisent une généralisation du théorème de Cauchy-Lipschitz qui nécessite un second membre one-sided Lipschitz.

Une importante contribution concernant l'analyse des EDOs discontinues a été fournie par Filippov [39], Utkin [69], Aubin et Cellina [7]. Dans [39] et [38], l'auteur reformule les problèmes d'EDOs discontinues comme des problèmes de contrôle ou comme des problèmes d'inclusion différentielle. En particulier, il prouve des résultats d'existence et d'unicité généraux. Dans le cas particulier où la surface de discontinuité est de codimension 1, l'auteur prouve un résultat d'unicité ne nécessitant pas de montrer la propriété de one-sided Lipschitz pour le second membre. Dans le cas des codimensions supérieures, le résultat proposé nécessite la non-oscillation des solutions et la non-tangence à la surface de discontinuité du champ de vecteur discontinu. Ces propriétés peuvent être complexes à obtenir en pratique. Un complément à cette théorie a été réalisé en étudiant les trajectoires possibles des solutions suivant la codimension et suivant la configuration du champ de vecteur discontinu [3], [4], [29], [30], [31], [32], [33] et [41]. En particulier, cela permet de définir une valeur du second membre discontinu sur les discontinuités.

Il existe des liens d'équivalence entre les systèmes de complémentarité, les inclusions différentielles et les systèmes de dynamique projetée, voir [13]. Les systèmes de dynamique projetée sont une autre manière de modéliser les EDOs discontinues qui, à notre connaissance, n'a pas été explorée dans le cadre de la cinétique. Cette théorie fournit des résultats d'existence et d'unicité de la solution [21], [20], [22] et [61] et une large gamme de méthodes numériques [1], [9], [8] et [59]. Dans [9] et [8], les auteurs proposent un algorithme qui permet de résoudre numériquement un système de dynamique projetée

sur une intersection d'ensembles convexes.

0.3 Descriptions des travaux de thèse

En géochimie, on s'intéresse à l'étude des interactions chimiques fluide-roche dans le sous-sol, qui ont des échelles de temps très diverses. Les cinétiques chimiques permettent de modéliser les réactions chimiques lentes. Contrairement aux réactions d'équilibre, les réactions de cinétique nécessitent la connaissance de la loi de vitesse de réaction qui lui est associée, là où les équilibres ne demandent que la connaissance de l'état final de la réaction appelé équilibre chimique. Cette thèse traite de la modélisation mathématique et de la simulation numérique des cinétiques chimiques.

Les modèles introduits dans cette thèse pour les systèmes de cinétique chimique prennent en compte les apparitions et disparitions de phase lors de processus chimiques. Ils étendent les résultats de [12] et [46] aux systèmes réactifs contenant plusieurs espèces en phase pure (minérale ou gazeuse) par réaction. Ces modèles permettent également de décrire des systèmes réactifs faisant intervenir des intermédiaires réactifs. La modélisation des systèmes de cinétique complexe, [6] et [40], est prise en charge par les modèles. De plus, les modèles proposés n'annulent pas nécessairement les vitesses des réactions chimiques lorsqu'une espèce disparaît.

On considère $\boldsymbol{c} = (\boldsymbol{c_a}, \boldsymbol{c_p})$ la concentration des espèces en phase aqueuse et en phase pure, le système d'équations relatif à la cinétique chimique est défini par

$$\frac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} = S\boldsymbol{r}, \;\; \boldsymbol{c}(0) = \boldsymbol{c_0} \geqslant 0, \tag{28}$$

où S est la matrice de stœchiométrie associée aux réactions de cinétiques, $\boldsymbol{c_0}$ la condition initiale et $\boldsymbol{r}$ représente le nouveau taux de réaction défini plus loin, voir (31).

Soit Q une base du noyau de S^T, si c est une solution de (28) alors

$$Q^T\boldsymbol{c} = Q^T\boldsymbol{c_0}. \tag{29}$$

Ce sont les lois de conservation intrinsèques au système chimique.

La loi de cinétique considérée est la loi d'action de masse cinétique définie pour une réaction j, quelconque, comme suit,

$$\tau_j := k_{f,j} \prod_{i\,:\,S_{i,j}<0} \Big(\frac{[E_i]}{K_i}\Big)^{|S_{i,j}|} - k_{b,j} \prod_{i\,:\,S_{i,j}>0} \Big(\frac{[E_i]}{K_i}\Big)^{|S_{i,j}|}, \tag{30}$$

où $[E_i]$ est l'activité d'une espèce, K_i est la constante d'équilibre d'une espèce, $k_{f,j}$ est la constante de vitesse du côté réactif et $k_{b,j}$ est la constante de vitesse du côté produit.

Remark 0.3.1. *La formule* (30) *fait apparaître les dépendances de k_r et k_p, définies en* (17)*, par rapport aux constantes d'équilibre K_i : $k_{r,j} = k_{f,j} \prod_{i:S_{i,j}<0} K_i^{-|S_{i,j}|}$ et $k_{p,j} = k_{b,j} \prod_{i:S_{i,j}>0} K_i^{-|S_{i,j}|}$.*

L'activité d'une espèce aqueuse est une fonction de la fraction molaire des espèces (on peut la réécrire en fonction de la concentration ou de la molalité). L'activité d'une espèce minérale en phase pure est égale à 1. Par conséquent, la loi de cinétique ne dépend pas des concentrations des phases pures. Comme montré dans [12], sans modification de la loi de vitesse de réaction, une espèce chimique en phase pure peut être amenée à prendre des valeurs de concentrations négatives.

Dans le premier chapitre, nous proposons un modèle discontinu de cinétique chimique permettant de traiter les apparitions et disparitions d'un nombre quelconque d'espèces minérales. Nous proposons une analyse détaillée d'un cas contenant une espèce réactive intermédiaire. Nous introduisons un modèle régularisé permettant de prouver un résultat d'existence et de positivité des solutions. Ce même modèle permet aussi de proposer une méthode numérique de résolution pour les cinétiques chimiques. Enfin, nous validons cette méthode numérique sur des exemples de systèmes réactifs.

Dans le deuxième chapitre, nous proposons une reformulation du modèle de cinétique chimique proposé dans le premier chapitre à l'aide de la théorie de Filippov. Nous réalisons une étude de l'existence et de la positivité des solutions de la formulation inclusion différentielle ainsi obtenue. Nous menons une étude de l'unicité et des trajectoires de la solution dans le cas des surfaces de discontinuité de codimension 1.

Dans le troisième chapitre, nous proposons un nouveau modèle de cinétique chimique de type système dynamique projetée. Nous réalisons une analyse de l'existence des solutions de ce modèle, des liens avec d'autres types de formulations et une méthode de résolution numérique adaptée. Nous illustrons les résultats numériques obtenus pour des systèmes de cinétique chimique.

0.3.1 Chapitre 1 : Modèles de cinétique chimique avec des phases pures

Ce travail fait l'objet d'un article en cours de rédaction en vue d'une soumission dans un journal. Les auteurs sont : Bastien Hamlat, Jocelyne Erhel, Anthony Michel et Thibault Faney.

Afin d'assurer la positivité de la concentration des espèces, on introduit une surface réactive dans la loi de cinétique. En particulier, une espèce disparue peut continuer à être consommée par le système réactif au plus à son taux de production. Par conséquent, la vitesse de réaction est définie par

$$r_j := h_j \tau_j, \tag{31}$$

avec $h_j \in [0, 1]$ un facteur permettant de limiter la vitesse de réaction lorsque la surface réactive est constante ou d'ordre 0. Nous avons étendu le concept de surface réactive d'ordre 0 développé dans [12] et [46] à des systèmes chimiques généraux comprenant des phases pures minérales et gazeuses ainsi qu'une phase aqueuse.

Précisons notre choix de modélisation pour $\boldsymbol{h}$:

Définition 0.3.1. *Les fonctions multivaluées U_j, $j = 1, \ldots, N_r$ sont définies dans $[0, +\infty)^{N_p}$ par :*

$$\begin{aligned} U_j(\boldsymbol{c}_p) &= \{1\} \text{ si } \forall k \in \mathcal{M}_j^-, c_k > 0, \text{ ou } \mathcal{M}_j^- = \varnothing, \\ &= [0,1] \text{ sinon.} \end{aligned}$$

Les fonctions multivaluées V_j, $j = 1, \ldots, N_r$ sont définies dans $[0, +\infty)^{N_p}$ par :

$$\begin{aligned} V_j(\boldsymbol{c}_p) &= \{1\} \text{ si } \forall k \in \mathcal{M}_j^+, c_k > 0, \text{ ou } \mathcal{M}_j^+ = \varnothing, \\ &= [0,1] \text{ sinon.} \end{aligned}$$

Le taux de réaction modifié est alors défini comme suit :

Définition 0.3.2. *Soit $T > 0$ et $\boldsymbol{c} \in (H^1(0,T))^{N_a+N_p}$ tel que $\boldsymbol{c}(t) \geqslant 0$, $0 \leqslant t \leqslant T$. Le taux de réaction modifié $\boldsymbol{r}$ est défini comme suit :*

$$r_j = u_j \tau_j^+(\boldsymbol{c}) - v_j \tau_j^-(\boldsymbol{c}),\ j = 1, \ldots, N_r, \tag{32}$$

où $u_j,\ v_j \in L^\infty(0,T)$ et $u_j(t) \in U_j(\boldsymbol{c}_p(t))$, $v_j(t) \in V_j(\boldsymbol{c}_p(t))$, $0 < t < T$.

Nous obtenons ainsi le modèle d'inclusion différentielle suivant :
trouver $\mathbf{c} \in (H^1(0,T))^{N_a+N_p}$, $\mathbf{u} \in (L^\infty(0,T))^{N_r}, \mathbf{v} \in (L^\infty(0,T))^{N_r}$ tel que

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = S\boldsymbol{r}(t, \mathbf{c}(t)),\ t \in (0,T), \\ \mathbf{c}(t) \geqslant 0, \\ \mathbf{c}(0) = \mathbf{c}_0, \\ u_j(t) \in U_j(\mathbf{c}_p(t)), v_j(t) \in V_j(\mathbf{c}_p(t)),\ t \in (0,T), \\ r_j = u_j \tau_j^+(\mathbf{c}_a) - v_j \tau_j^-(\mathbf{c}_a),\ j = 1, \ldots, N_r, \end{cases} \tag{33}$$

où $\mathbf{c}_0 \geqslant 0$ est la condition initiale, U_j et V_j sont introduits en Définition 0.3.1.

On peut noter que les lois de conservation (29) sont respectées par ce modèle. Les objectifs sont alors de caractériser la surface réactive u_j et v_j, de prouver l'existence d'une solution et de construire un algorithme de résolution.

Une première contribution du chapitre est la construction d'un modèle régularisé. Nous appliquons une fonction de Heaviside régularisée H_ϵ à chaque phase pure, puis une fonction mélange aux phases pures impliquées dans une réaction. Nous définissons ainsi des surfaces réactives régularisées $\boldsymbol{u_\epsilon}$ et $\boldsymbol{v_\epsilon}$, ainsi qu'une vitesse de réaction régularisée $\boldsymbol{r_\epsilon}$:

Définition 0.3.3. *Le taux de réaction régularisé est défini comme suit :*

$$r_{j,\epsilon}(\boldsymbol{c}) = u_{j,\epsilon}(\boldsymbol{c}) \tau_j^+(\boldsymbol{c}) - v_{j,\epsilon}(\boldsymbol{c}) \tau_j^-(\boldsymbol{c}), \tag{34}$$

où $u_{j,\epsilon}$ and $v_{j,\epsilon}$ sont définis par

$$\begin{cases} \text{si } \mathcal{M}_j^- = \varnothing, \text{ alors } u_{j,\epsilon}(\boldsymbol{c}) = 1, \text{ sinon } u_{j,\epsilon}(\boldsymbol{c}) = \min\limits_{k \in \mathcal{M}_j^-} (H_\epsilon(c_k)), \\ \text{si } \mathcal{M}_j^+ = \varnothing, \text{ alors } v_{j,\epsilon}(\boldsymbol{c}) = 1, \text{ sinon } v_{j,\epsilon}(\boldsymbol{c}) = \min\limits_{k \in \mathcal{M}_j^+} (H_\epsilon(c_k)). \end{cases} \tag{35}$$

Nous définissons le système cinétique réactif régularisé comme suit :

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = S\boldsymbol{r}_{\boldsymbol{\epsilon}}(\mathbf{c}(t)),\ t \geqslant 0, \\ \mathbf{c}(0) = \mathbf{c}_0. \end{cases} \tag{36}$$

Ce système (36) admet une unique solution c^ϵ.

Théorème 0.3.1. *On suppose qu'il existe un vecteur $\boldsymbol{q}$ tel que $\boldsymbol{q} > 0$ et $S^T\boldsymbol{q} = 0$. le système (36) admet une unique solution $\boldsymbol{c}_\epsilon$ dans $[0, +\infty[$ qui satisfait*

$$0 \leqslant \boldsymbol{c}_\epsilon \leqslant \boldsymbol{q}^T \frac{\boldsymbol{c}_0}{\min\limits_k(q_k)}.$$

Une deuxième contribution du chapitre est un résultat d'existence des solutions faibles de (33) grâce à une analyse de convergence de (36).

Théorème 0.3.2. *On suppose qu'il existe un vecteur $\boldsymbol{q}$ tel que $\boldsymbol{q} > 0$ et $S^T\boldsymbol{q} = 0$. Soit $T > 0$. La suite $\boldsymbol{c}_\epsilon$, solution de (36), converge (à une sous-suite près) fortement dans $C^0(0, T, \mathbb{R}^{(N_a+N_p)})$. De plus, toute limite $\boldsymbol{c}$ vérifie $0 \leqslant \boldsymbol{c} \leqslant \boldsymbol{q}^T \dfrac{\boldsymbol{c}_0}{\min\limits_k(q_k)}$.*

Pour tout $j = 1, \ldots, N_r$, les contrôles régularisés $u_{j,\epsilon}(\boldsymbol{c}_\epsilon)$ et $v_{j,\epsilon}(\boldsymbol{c}_\epsilon)$ admettent des limites au sens faible- dans $L^\infty(0, T, \mathbb{R}^{N_r})$ notés u_j et v_j respectivement. De plus, pour tout $j = 1, \ldots, N_r$ et $0 < t < T$, les limites $\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}$ satisfont $u_j(t) \in U_j(\boldsymbol{c}_p(t))$ et $v_j(t) \in V_j(\boldsymbol{c}_p(t))$. Enfin $\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}$ sont des solutions faibles de (33).*

Une troisième contribution du chapitre est l'analyse d'un cas de cinétique chimique : $M_1 \rightleftharpoons M_2$ et $M_2 \rightleftharpoons M_3$. Dans ce contexte, le système discontinu est défini comme suit :

$$\begin{cases} \dfrac{dc_1}{dt} = -r_1, \\ \dfrac{dc_2}{dt} = r_1 - r_2, \\ \dfrac{dc_3}{dt} = r_2, \\ c(0) = c_0, \\ c \geqslant 0, \end{cases} \tag{37}$$

avec $r_1 = u_1\tau_1^+ - v_1\tau_1^-$ et $r_2 = u_2\tau_2^+ - v_2\tau_2^-$.

Après une analyse des différents contextes possibles, on obtient la formulation du modèle discontinu limite. Le taux de réaction discontinu dans ce cas est défini comme suit :

$$\begin{cases} r_1 = H(c_1)\tau_1^+ - H(c_2)\tau_1^- - \delta(c_2)H(c_3)\tau_2^-, \\ r_2 = H(c_2)\tau_2^+ + \delta(c_2)H(c_1)\tau_1^+ - H(c_3)\tau_2^-. \end{cases} \tag{38}$$

avec $\delta : x \to \begin{cases} 0 \text{ si } x \in \mathbb{R}^* \\ 1 \text{ si } x = 0 \end{cases}$ et H la fonction de Heaviside.

La dernière contribution de ce chapitre est la proposition d'un schéma numérique basé sur la formulation régularisée (36), qu'on illustre sur plusieurs cas chimiques : $A \rightleftharpoons M$, $M_1 \rightleftharpoons M_2 \rightleftharpoons M_3$, $A_1 \rightleftharpoons M_2 \rightleftharpoons A_3$ et d'autres cas plus complexes.

0.3.2 Chapitre 2 : Analyse d'un modèle de type EDOs discontinues pour la cinétique chimique multiphasique

Dans ce chapitre, nous utilisons le cadre des équations différentielles discontinues de la théorie de Filippov [39] pour analyser le système (33).

Soit i l'indice d'un minéral, on note $\Sigma_i = \{c_i = 0\}$ une surface de discontinuité. Ces surfaces séparent l'espace des phases pures en des domaines G_k qui correspondent à des portions de l'espace avec un signe constant pour les concentrations des minéraux. Chaque domaine G_k est caractérisé par un vecteur $\boldsymbol{e^k}$: pour tout $i \in \{1, \cdots, N_p\}$, $e_i^k = \begin{cases} 1, & \text{si } c_i > 0, \\ 0, & \text{si } c_i < 0. \end{cases}$
On note G l'union disjointe des G_k et Σ l'union des Σ_i.

Pour construire le modèle EDOs discontinues, l'idée est de prolonger le modèle (33) dans les domaines G_k où au moins une concentration des minéraux est négative. Dans notre modèle, les limiteurs de réaction valent 0 dans le domaine G_k si au moins un des minéraux à concentration négative est consommé par la réaction j.

On définit cette extension comme suit, $\boldsymbol{r^k} : \overline{G_k} \mapsto \mathbb{R}^{N_r}$ tel que

$$r_j^k(\boldsymbol{c}) = \Big(\prod_{i \in \mathcal{M}_j^-} e_i^k \Big) \tau_j^+(\boldsymbol{c}) - \Big(\prod_{i \in \mathcal{M}_j^+} e_i^k \Big) \tau_j^-(\boldsymbol{c}), \ \forall k \in \{1, \cdots, 2^{N_p}\}, \ \forall j \in \{1, \cdots, N_r\}. \quad (39)$$

Afin de définir le modèle d'inclusions différentielles permettant l'utilisation de la théorie de Filippov, on introduit les notations suivantes :
$\mathcal{G}(\boldsymbol{c})$ est l'ensemble des indices des domaines G_k adhérent à $\boldsymbol{c}$.
Le taux de réaction multivalué utilisant la convexification de Filippov est défini comme suit

$$\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}) = \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k \boldsymbol{r^k}(\boldsymbol{c}), \ \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c}), \quad (40)$$

avec

$$\Lambda(\boldsymbol{c}) = \{\boldsymbol{\lambda} \in [0,1]^{\text{card } \mathcal{G}(\boldsymbol{c})}, \ \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k = 1\}. \quad (41)$$

Il en résulte le système d'inclusion différentielle suivant :

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} \in \big\{ S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}), \ \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c}) \big\}, \\ \boldsymbol{c}(0) = \boldsymbol{c}_0 \geqslant 0. \end{cases} \quad (42)$$

Dans la suite on note $\boldsymbol{F}(\boldsymbol{c}) = \big\{ \boldsymbol{f}(\boldsymbol{\lambda}, \boldsymbol{c}), \ \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c}) \big\}$ avec $\boldsymbol{f}(\boldsymbol{\lambda}, \boldsymbol{c}) = S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c})$, le second membre multivalué au sens du convexifié de Filippov.

Le premier résultat de ce chapitre assure que toute fonction absolument continue solution de (42) est à valeur positive ou nulle.

Théorème 0.3.3. *Soit* $\boldsymbol{c} : [0,T] \longmapsto \mathbb{R}^{N_a+N_p}$ *une solution de* (42). *Alors* $\forall t \in [0,T]$, $\boldsymbol{c}(t) \in C_0 = \{\boldsymbol{c^*} \in (\mathbb{R}_+)^{N_a+N_p}, \ Q^T\boldsymbol{c^*} = Q^T\boldsymbol{c}_0\}$.

De plus, grâce à la théorie de Filippov [39], nous montrons l'existence d'une solution de (42).

Théorème 0.3.4. *Sous l'hypothèse qu'il existe* $\boldsymbol{q} > 0$ *tel que* $\boldsymbol{q} \in \mathrm{Ker}(S^T)$*, pour tout* $\boldsymbol{c_0} \in (\mathbb{R}_+)^{N_a+N_p}$*, il existe une solution* $\boldsymbol{c} : [0, +\infty[\to C_0$ *de* (42)*.*

Le deuxième résultat concerne l'unicité de la solution de (42) ainsi que l'étude des trajectoires dans le cas où un seul minéral disparaît. Pour ce faire, l'idée est d'effectuer une classification et une étude de toutes les configurations possibles de champs signés afin d'en déduire les trajectoires des solutions de (42). En particulier, l'étude montre que le cas répulsif n'est pas possible en cinétique chimique.

Dans le contexte de l'étude des discontinuités de codimension 1 on définit $G_1 = \{(\boldsymbol{c_a}, \boldsymbol{c_p}),\ \boldsymbol{c_a} \geqslant 0,\ \boldsymbol{c_p} > 0\}$ et $G_2 = \{(\boldsymbol{c_a}, \boldsymbol{c_p}),\ \boldsymbol{c_a} \geqslant 0,\ c_{p,1} < 0,\ \boldsymbol{c_{p,i}} > 0,\ i \geqslant 2\}$ les deux ensembles autour de la surface de discontinuité $\Sigma_1^+ = \{(\boldsymbol{c_a}, \boldsymbol{c_p}),\ \boldsymbol{c_a} \geqslant 0,\ c_{p,1} = 0,\ c_{p,i} > 0,\ i \geqslant 2\}$. De la même manière, $\boldsymbol{f^1}$ et $\boldsymbol{f^2}$ sont les second-membres discontinus de (42) dans les domaines $\overline{G_1}$ et $\overline{G_2}$.

Le second-membre multivalué $\boldsymbol{F}$ de (42) est défini comme suit

$$\boldsymbol{F}(\boldsymbol{c}) = \begin{cases} \{\boldsymbol{f^1}(\boldsymbol{c})\}, & \boldsymbol{c} \in G_1, \\ \{\boldsymbol{f^2}(\boldsymbol{c})\}, & \boldsymbol{c} \in G_2, \\ \{\lambda \boldsymbol{f^1}(\boldsymbol{c}) + (1-\lambda)\boldsymbol{f^1}(\boldsymbol{c}),\ \lambda \in [0,1]\}, & \boldsymbol{c} \in \Sigma_1^+. \end{cases} \tag{43}$$

Notre étude des configurations des champs de vecteur de cinétique chimique nous conduit aux trois cas possibles décrits ci-dessous :
Soit $\boldsymbol{c} \in \Sigma_1^+$,

— Attractif : $[f_1^1(\boldsymbol{c}) < 0$ (et $f_1^2(\boldsymbol{c}) \geqslant 0)]$ ou $[f_1^1(\boldsymbol{c}) = 0$ et $f_1^2(\boldsymbol{c}) > 0]$.
— Transverse : $f_1^1(\boldsymbol{c}) > 0$ (et $f_1^2(\boldsymbol{c}) > 0$).
— Tangent : $f_1^1(\boldsymbol{c}) = 0$ et $f_1^2(\boldsymbol{c}) = 0$.

Dans les cas attractif et transverse on prouve l'unicité de la solution de (42) en s'appuyant sur des résultats de la théorie de Filippov dans le contexte des surfaces de discontinuité de codimension 1.

0.3.3 Chapitre 3 : Modèle avec dynamique projetée pour les réactions de dissolution et de précipitation des minéraux en géochimie

Ce travail fait l'objet d'un article en cours de rédaction en vue d'une soumission dans un journal. Les auteurs sont : Jocelyne Erhel, Bastien Hamlat et Tangi Migot.

Dans ce travail, réalisé en collaboration avec Jocelyne Erhel et Tangi Migot, nous proposons un autre modèle de cinétique chimique, une analyse d'existence des solutions et une méthode numérique de résolution. La première contribution de ce chapitre est de proposer un modèle pour lequel la dynamique de chaque espèce est limitée par un facteur $\lambda_i \in [0,1]$ de sorte que $\dfrac{\mathrm{d}c_i}{\mathrm{d}t} = \lambda_i (S\tau)_i$. On note $D(\boldsymbol{c})$ l'ensemble qui vérifie cette contrainte, défini ci-dessous,

$$D(c) := \{\mathrm{diag}(S\boldsymbol{\tau}(\boldsymbol{c}))\boldsymbol{\lambda} \text{ with } 0 \leqslant \boldsymbol{\lambda} \leqslant 1\}.$$

De plus, les concentrations des espèces chimiques doivent respecter les lois de conservation et la positivité. On note C_0 l'ensemble associé à ces contraintes, défini comme suit,

$$C_0 := \{\boldsymbol{c} \geqslant 0 \ : \ Q^T(\boldsymbol{c} - \boldsymbol{c_0}) = 0\}.$$

Cette approche aboutit à un système projectif de façon à garantir la positivité de $\boldsymbol{c}$, à respecter les lois de conservation et à maintenir la vitesse des concentrations proportionnelle à la vitesse non limitée. Le modèle proposé et étudié dans ce chapitre est le suivant

$$\begin{cases} \dot{c}(t) = \text{proj}\left(D(\boldsymbol{c}(t)) \cap \mathcal{T}(C_0, \boldsymbol{c}(t)); S\boldsymbol{\tau}(\boldsymbol{c}(\boldsymbol{t}))\right), \\ \boldsymbol{c}(t) \in C_0, \\ \boldsymbol{c}(0) = \boldsymbol{c_0}, \end{cases} \tag{44}$$

où $\mathcal{T}(C_0; \boldsymbol{c})$ est le cône tangent de C_0 en $\boldsymbol{c}$. La projection d'un vecteur v sur un ensemble E est définie par $\text{proj}(E; v) := \{w \in E : \|v - w\|_2 = \inf_{z \in E} \|v - z\|_2\}$.

La deuxième contribution de ce chapitre est un lien entre le modèle (44) et d'autres modèles : complémentarité dans le cas $A \rightleftharpoons M$, discontinu (33) dans le cas $M_1 \rightleftharpoons M_2$ et $M_2 \rightleftharpoons M_3$ et dynamique projeté sans la contrainte $\dot{\boldsymbol{c}} \in D(\boldsymbol{c})$ dans le cas $M_1 \rightleftharpoons M_2$ et $M_2 \rightleftharpoons M_3$.

La troisième contribution de ce chapitre est la démonstration de l'existence des solutions de (44). Pour cela, nous proposons une étude d'un schéma discrétisé de (44) ainsi qu'une analyse de convergence. Nous utilisons un schéma explicite en temps transformant (44) en un problème non linéaire à chaque pas de temps. Soit $h = \frac{t_f}{N_h}$ où N_h est le nombre de pas de temps et t_f le temps final. Pour tout $k = 0, \cdots, N_h - 1$, nous définissons $\boldsymbol{c^{k+1}}$ par

$$\begin{cases} \boldsymbol{c^{k+1}} \in K_k, \\ \boldsymbol{c^{k+1}} = \text{proj}(K_k; \boldsymbol{c^k} + h\boldsymbol{S\tau}(\boldsymbol{c^k})), \end{cases} \tag{45}$$

où $K_k = C_0 \cap (\boldsymbol{c^k} + hD(\boldsymbol{c^k}))$. Il est à noter que grâce aux propriétés de K_k, $\boldsymbol{c^{k+1}}$ existe et est unique.

Nous prouvons l'existence d'une solution de (44) en construisant une fonction $\boldsymbol{c_h}$, définie pour $t \in [0, t_f]$ et $k = \left\lfloor \frac{t}{h} \right\rfloor$ par

$$\boldsymbol{c_h}(t) = \boldsymbol{c_h^k} + (\boldsymbol{c_h^{k+1}} - \boldsymbol{c_h^k})\frac{t - kh}{h}. \tag{46}$$

On a le résultat suivant

Théorème 0.3.5. *$\boldsymbol{c_h}$ converge uniformément vers $\boldsymbol{c}$ solution de* (44).

La preuve repose sur des résultats d'analyse classique comme le théorème d'Arzela-Ascoli et le lemme de Mazur. La dernière contribution de ce chapitre concerne la simulation numérique. Nous utilisons, l'algorithme itératif avec double projection pour calculer $\boldsymbol{c^{k+1}}$ pour une tolérance $\epsilon > 0$:

$$
\begin{aligned}
&\boldsymbol{c_0^{k+1}} = \boldsymbol{c^k} + hS\boldsymbol{\tau}(\boldsymbol{c^k}), \boldsymbol{e_0} = \boldsymbol{e_{-1}} = 0,\\
&j = 0,\\
&\textbf{Répéter en incrémentant } j \textbf{ :}\\
&\boldsymbol{c_{2j+1}^{k+1}} = \mathrm{proj}(C_0; \boldsymbol{c_{2j}^{k+1}} + \boldsymbol{e_{2j-1}}),\\
&\boldsymbol{e_{2j+1}} = \boldsymbol{c_{2j}^{k+1}} + \boldsymbol{e_{2j-1}} - \boldsymbol{c_{2j+1}^{k+1}},\\
&\boldsymbol{c_{2j+2}^{k+1}} = \boldsymbol{c^k} + h\mathrm{proj}\left(D(\boldsymbol{c^k}); \frac{\boldsymbol{c_{2j+1}^{k+1}} + \boldsymbol{e_{2j}} - \boldsymbol{c^k}}{h}\right),\\
&\boldsymbol{e_{2j+2}} = \boldsymbol{c_{2j+1}^{k+1}} + \boldsymbol{e_{2j}} - \boldsymbol{c_{2j+2}^{k+1}},\\
&\textbf{jusqu'à } \|\boldsymbol{c_{2j}^{k+1}} - \boldsymbol{c_{2j-1}^{k+1}}\| \leqslant \epsilon
\end{aligned}
\tag{47}
$$

alors $\boldsymbol{c^{k+1}} = \boldsymbol{c_{2j}^{k+1}}$.

Cet algorithme est programmé en Matlab afin de résoudre (44) numériquement. Une validation de la méthode numérique est effectuée sur des cas tests.

Chapter 1

Kinetic models with pure phases

This work corresponds to a preprint in view of a submission in a journal. The authors are: Bastien Hamlat, Jocelyne Erhel, Anthony Michel and Thibault Faney.

1.1 Introduction

In this paper, we are interested in reactive transport modeling. This kind of models are a coupled ODEs-PDEs, where the PDEs describe the fluid transport and the ODEs describe the kinetic reaction, for reactive transport models including minerals see [49], [65], [70], [71] and [72]. Algebraic constraints are also included in order to take into account equilibrium reactions in the model. In the present work, we focus on ODEs models for chemical kinetic systems.

In chemical kinetics, the speed of a reaction is described by the reaction rate. The reaction rate is modeled by the law of mass action [18], a nonlinear kinetic law, see Section 1.2.4. As proved in several papers, [12] and [46], this law is not mathematically adapted to take into account the disappearance of a mineral. In order to handle this, we consider kinetic-controlled models [2], [12], [40], [46], [49], [51], [52] and [72]. A first possibility is to consider a discontinuous reaction rate, see Section 1.3.1 for an introduction of this model and for more details [12] and [40]. A second solution is to consider a set-va lued model, see Section 1.3.2 for an introduction and for more details [2], [49] and [72]. A third possibility is to consider complementarity conditions in order to change the value of the reaction rate depending on the presence of the mineral, see Section 1.3.3 for an introduction and [46] for more details. All of these models allow to consider one mineral in dissolution or precipitation in a kinetic reaction.

The aim of this paper is to investigate the case of several minerals in a kinetic reaction or a mineral involved in several reactions. The physical motivation is to consider complex chemical systems as described in [6]. The main difficulty is to propose a kinetic model which allows to maintain the reaction process when one of the minerals reactant disappear. In this paper, we generalize the set-valued formulation for kinetic problems containing several minerals in each reaction, see [40] and [65] for example of reaction containing several minerals. This generalization is obtained by introducing two limiter set-valued functions for each reaction (one for the reactant side and another for the product side). Each of this limiters can depend on several minerals, i.e. each reactant limiter depends on all the reactants in its reaction. We follow the idea of [72] by proposing a

regularized model to prove the existence of solutions of the set-valued model. Moreover, one novelty of this work is to use the regularized formulation for numerical experiments. Keeping in mind that the aim is to couple the kinetic model in a fully implicit reactive transport solver [52] and [56].

This paper is divided into five main parts. The first one is a general introduction to chemical kinetics. The second one gathers the state of the art in the case of one mineral in one reaction. In the third part, we present our set-valued model, its regularization and associated convergence and existence results. In the fourth part, we investigate the case $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$. The last part concerns numerical experiments: simulation, convergence and an example of non-uniqueness.

1.2 Mathematical model of kinetic reactions

We consider a geochemical system with an aqueous phase and several mineral phases. The aqueous phase is composed of water and several diluted species whereas a mineral phase contains only one species. We do not consider source or sink terms. We assume that the pressure and temperature are constant during the reactions.

Kinetic reactions in this system are governed by reaction rates and conservation laws. Reactions are considered reversible, with forward and backward directions. The number of reactions is N_r and each reaction is noted R_j, $j = 1, \dots, N_r$. The number of aqueous species is $N_a > 0$ and each aqueous species is noted A_i, $i = 1, \dots, N_a$. The number of minerals is $N_p \geqslant 0$ and if $N_p > 0$, each mineral is noted M_k, $k = N_a + 1, \dots, N_a + N_p$. The stoichiometric matrix $S \in \mathbb{Z}^{(N_a+N_p)\times N_r}$ contains the coefficients of reactions.

We set an arbitrary forward direction of a reversible reaction R_j. Species on the left are called reactants and are associated to negative stoichiometric coefficients, whereas species on the right, called products, correspond to positive coefficients [56]. We denote by $\mathcal{A}_j^-$ and $\mathcal{M}_j^-$ the sets of aqueous and mineral reactants in reaction R_j, with negative stoichiometric coefficients $s_{ij} < 0$; by $\mathcal{A}_j^+$ and $\mathcal{M}_j^+$ the sets of aqueous and mineral products, with positive coefficients $s_{ij} > 0$.

A reaction R_j is written

$$\sum_{i \in \mathcal{A}_j^-} |s_{ij}| A_i + \sum_{k \in \mathcal{M}_j^-} |s_{kj}| M_k \rightleftarrows \sum_{i \in \mathcal{A}_j^+} |s_{ij}| A_i + \sum_{k \in \mathcal{M}_j^+} |s_{kj}| M_k. \tag{1.1}$$

1.2.1 Examples

We provide here two examples of such reactions in geology.

In the first example, calcite $CaCO_3$ reacts with calcium oxide CaO which reacts with calcium hydroxide $Ca(OH)_2$:

$$\begin{aligned} R_1 :&\quad CaCO_3 \rightleftarrows CaO + CO_2, \\ R_2 :&\quad CaO + H_20 \rightleftarrows Ca(OH)_2. \end{aligned}$$

The second example still concerns calcite, which can react with dolomite $CaMg(CO_3)_2$ and siderite $FeCO_3$ when the solution contains magnesium and irons:

$$R_1 : \ 2CaCO_3 + Mg^{++} \rightleftarrows CaMg(CO_3)_2 + Ca^{++},$$
$$R_2 : \ 3CaCO_3 + Mg^{++} + Fe^{++} \rightleftarrows CaMg(CO_3)_2 + FeCO_3 + 2Ca^{++}.$$

Finally, it can be noted that calcite is a form of calcium carbonate which can interact with another form, aragonite, under some conditions.

In the domain of electrochemistry, reactions between solute and minerals occur in batteries.

1.2.2 Quantities of species and reaction rates

The unknowns in such a geochemical system are the number of moles of each species, grouped into a vector of functions $\mathbf{c} = (c_i),\ i = 1, \ldots, N_a + N_p$ of the time variable t with $t \geqslant 0$. We denote by $\mathbf{c}_a = (c_i), i = 1, \ldots, N_a$ the vector of quantities for aqueous species, and by $\mathbf{c}_p = (c_k), k = N_a + 1, \ldots, N_a + N_p$ the vector of mineral quantities. The quantities c should satisfy a positivity constraint $\mathbf{c}(t) \geqslant 0,\ t \geqslant 0$. We also require that $\mathbf{c} \in (H^1(0,T))^{N_a+N_p}$ with $T > 0$.

In this paper, we propose a model of differential equations, which satisfy mass conservation laws and various invariants. The right-hand side uses reaction rates $\boldsymbol{r} = (r_j)$, functions in $L^2(0,T)$, based on mass action laws. Our objective is to define a model of reaction rates such that the differential system has a solution satisfying the requirements given above.

1.2.3 Conservation laws

Mass conservation laws are grouped into a system of differential equations:

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = S\mathbf{r}(t, \mathbf{c}(t)), t \in (0,T) \\ \mathbf{c}(t) \geqslant 0, \\ \mathbf{c}(0) = \mathbf{c}_0, \end{cases} \tag{1.2}$$

where $\mathbf{c}_0 \geqslant 0$ is a given initial condition, and reaction rates $\mathbf{r}$ are discontinuous functions, defined thanks to mass action laws and set-valued functions. We define a weak solution of system (1.2).

Definition 1.2.1. *Let $T > 0$, let $\boldsymbol{c} \in H^1(0,T)^{N_a+N_p}$. It is a weak solution of system (1.2) if $\boldsymbol{c} \geqslant 0$ and $\forall t \in [0,T]$,*

$$c_i(t) = c_{0,i} + \int_0^t (Sr)_i(s, \boldsymbol{c}(s))ds,\ i = 1, \ldots, N_a + N_p.$$

It can be noted that this system implies invariants, which correspond to various balance equations. Indeed, let Q be a basis of the kernel of S^T, such that $Q^T S = S^T Q = 0$. Then, any solution of (1.2) satisfies $\forall t \geqslant 0, Q^T\mathbf{c}(t) = Q^T\mathbf{c}_0$. Functions $Q^T\mathbf{c}$ are thus constant in time and are conservative variables. For example, if some species are ions, then one conservative variable is the charge balance.

If a conservative variable $\mathbf{q}^T\mathbf{c}$ is defined by a positive vector $\mathbf{q} > 0$, then quantities of species are bounded:

Lemma 1.2.1. *Let $\boldsymbol{q}$ be a vector such that $\boldsymbol{q} > 0$. If $\boldsymbol{c}$ satisfies $\boldsymbol{c} \geqslant 0$ and $\boldsymbol{q}^T \boldsymbol{c} = \boldsymbol{q}^T \boldsymbol{c}_0$ then $\boldsymbol{c} \leqslant C$ with $C = \dfrac{1}{\min\limits_k(q_k)} \boldsymbol{q}^T \boldsymbol{c}_0$.*

Proof. Let $k \in \{1, \ldots, N_a + N_p\}$. First, $\mathbf{q}^T\mathbf{c} = q_k c_k + \sum\limits_{i \neq k} q_i c_i \geqslant q_k c_k$ since $c_i \geqslant 0$ and $q_i > 0$. Then $q_k c_k \leqslant \mathbf{q}^T \mathbf{c}_0$ and $c_k \leqslant \dfrac{1}{q_k} \mathbf{q}^T \mathbf{c}_0 \leqslant C$. □

Physically, this result prohibits the chemical system from creating matter. In other words, the total mass of each atom is an invariant of the chemical system.

1.2.4 Maximal reaction rates

Reactions R_j involve a forward reaction, with reactants and products, and a backward reaction, where products become reactants and vice versa. Each direction has its own reaction rate, and the global rate is the difference between both. We call it the maximal reaction rate.

When a reaction involves only aqueous species, its rate is given by a mass action law, which is a function of species' activities [18]. When a pure phase exists, its activity is equal to 1; we extend it to 1 when the quantity of pure phase is zero. Various models of aqueous activities exist in the literature, see [16] for example. Throughout the paper, we make the following assumptions.

Assumption 1.2.1. *The activity a_i of a solute A_i is a C^1 function of $\boldsymbol{c}_a$ which satisfies:*

$$\boldsymbol{c}_a \geqslant 0 \Rightarrow \boldsymbol{a}(\boldsymbol{c}_a) \geqslant 0$$

$$c_{a,i} = 0 \Rightarrow a_i(\boldsymbol{c}_a) = 0.$$

The forward reaction rate of reaction R_j is given by

$$\tau_{f,j} = k_{f,j} \prod_{i \in \mathcal{A}_j^-} \left(\frac{a_i}{K_i}\right)^{|s_{ij}|} \prod_{k \in \mathcal{M}_j^-} \left(\frac{1}{K_k}\right)^{|s_{kj}|}$$

and its backward reaction rate is

$$\tau_{b,j} = k_{b,j} \prod_{i \in \mathcal{A}_j^+} \left(\frac{a_i}{K_i}\right)^{|s_{ij}|} \prod_{k \in \mathcal{M}_j^+} \left(\frac{1}{K_k}\right)^{|s_{kj}|}$$

If $\mathcal{A}_j^-$ is empty, the product $\prod\limits_{i \in \mathcal{A}_j^-} \left(\dfrac{a_i}{K_i}\right)^{|s_{ij}|}$ is equal to 1. The same property applies to other subsets. The maximal reaction rate is the difference between forward and backward reaction rates:

$$\tau_j = \tau_{f,j} - \tau_{b,j}.$$

Finally, we will use the following formula, where constants are grouped together:

$$\tau_j(\mathbf{c}_a) = \alpha_j \prod_{i \in \mathcal{A}_j^-} a_i(\mathbf{c}_a)^{|s_{ij}|} - \beta_j \prod_{i \in \mathcal{A}_j^+} a_i(\mathbf{c}_a)^{|s_{ij}|}, j = 1, \ldots, N_r. \tag{1.3}$$

where α_j and β_j are strictly positive constants of reactions.

Thermodynamic equilibrium of reaction R_j is obtained when forward and backward reactions are balanced, in other words when the reaction rate is zero: $\tau_j = 0$. It corresponds to the mass action law for chemistry at equilibrium.

1.2.5 Continuous model

We consider the continuous differential system with the vector $\boldsymbol{\tau}$ of maximal reaction rates:

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = S\boldsymbol{\tau}(\mathbf{c}_a(t)), t \geqslant 0, \\ \mathbf{c}(0) = \mathbf{c}_0. \end{cases} \tag{1.4}$$

Proposition 1.2.1. *The geochemical system (1.4) has a unique solution satisfying* $\boldsymbol{c} \in (H^1(0,+\infty))^{N_a+N_p}$.

Proof. Thanks to Assumption 1.2.1, the maximal rate $\boldsymbol{\tau}$ is C^1. Since the right-hand side is regular, we can apply theorem of Cauchy-Lipschitz so that the system has a unique solution in $(0,+\infty)$. □

In order to study the non-negativity of quantities c_i, we introduce the notion of essential non negativity [18].

Definition 1.2.2. *Let f be a function from $\mathbb{R}^{N_a}$ to $\mathbb{R}^{N_a}$. It is essentially non-negative if, for any $k = 1, \ldots, N_a$,*

$$c_i \geqslant 0, i = 1, \ldots, N_a \text{ and } c_k = 0 \Rightarrow f_k(c) \geqslant 0.$$

Since the reaction rates τ_j depend only on the aqueous species, we can define a function f from $\mathbb{R}^{N_a}$ to $\mathbb{R}^{N_a}$ by

$$f_k(\mathbf{c}_a) = (S\tau(\mathbf{c}_a))_k, \ k = 1, \ldots, N_a. \tag{1.5}$$

Proposition 1.2.2. *The function f defined by (1.5) is locally Lipschitz continuous and essentially non negative.*

Proof. Since $\boldsymbol{\tau}$ is regular, f is at least locally Lipschitz continuous.

Let $\mathbf{c}_a \geqslant 0$ with $c_k = 0$ for some k. Then $\mathbf{a}(\mathbf{c}_a) \geqslant 0$ and $a_k(c_k) = 0$.

If $s_{kj} > 0$ then $A_k \in \mathcal{A}_j^+$ and $\tau_j(\mathbf{c}_a) \geqslant 0$ thus $s_{kj}\tau_j \geqslant 0$.

If $s_{kj} < 0$ then $A_k \in \mathcal{A}_j^-$ and $\tau_j(\mathbf{c}_a) \leqslant 0$ thus $s_{kj}\tau_j \geqslant 0$.

Finally, $f_k = \sum_{j:\, s_{kj} \neq 0} s_{kj}\tau_j(\mathbf{c}_a) \geqslant 0$. □

Thanks to this property, we can prove that the quantities of all aqueous species remain positive.

Proposition 1.2.3. *If $\boldsymbol{c}_0 \geqslant 0$ then the unique solution $\boldsymbol{c}$ of system (1.4) satisfies $\boldsymbol{c}_a \geqslant 0$.*

Proof. Let us consider the system of ODE with only the first N_a equations, in other words with the function f, which is essentially non negative and locally Lipschitz continuous by Proposition 1.2.2. By using Theorem 1 in [18], we get the result wanted. □

On the other hand, denoting by $\mathbf{c}$ the solution of (1.4) it may exist a mineral M_k, $N_a + 1 \leqslant k \leqslant N_a + N_p$ and a time t^* such that $c_k(t^*) = 0$ and $\forall t > t^*$, $c_k(t) < 0$. Models of reaction rates r aim at limiting the reaction so that the solution of (1.2) is non-negative. We first consider a specific case, well studied in the literature.

1.3 State of the art: one mineral in one reaction

Here, we consider the particular case of two aqueous species A_1 and A_2 which precipitate into a mineral M:

$$s_1 A_1 + s_2 A_2 \rightleftarrows M$$

with $s_1 > 0, s_2 > 0$. An example is the precipitation-dissolution of silver chloride: $\mathrm{Ag}^+_{(w)} + \mathrm{Cl}^-_{(w)} \rightleftharpoons \mathrm{AgCl}_{(s)}$. Such kinetic models have been studied by several authors [2, 12, 46]. Since the system of three equations has two invariants, it is sufficient to consider the equation for c_M:

$$\frac{\mathrm{d}c_M}{\mathrm{d}t} = r.$$

The maximal reaction rate given by (1.3) is

$$\tau = \alpha a_1^{s_1} a_2^{s_2} - \beta,$$

where a_1, a_2 are activities of aqueous species A_1, A_2. It seems natural to choose $r = \tau$ as long as c_M is not fully dissolved ($c_M > 0$). Different choices of r appear in the literature when $c_M = 0$. With a discontinuous model, r is such that the reaction stops as soon as the mineral is fully dissolved [12]. With a set-valued model, r belongs to a given set [46]; moreover, the value at $c_M = 0$ can be specified [2]. A third model is based on a complementarity problem [46]. Under some assumptions, the three models are equivalent in the sense that they have the same unique weak solution [46]. Another model defines a projected dynamical system, which limits the derivative of each mineral species [34].

1.3.1 Discontinuous model

The discontinuous model implies $r = 0$ after dissolution [12]:

$$\begin{cases} \text{if } c_M > 0 \text{ or } \tau \geqslant 0 \text{ then } r = \tau, \\ \text{if } c_M = 0 \text{ and } \tau \leqslant 0 \text{ then } r = 0. \end{cases} \tag{1.6}$$

Let us introduce the positive part $\tau^+(\mathbf{c}_a)$ and the negative part $\tau^-(\mathbf{c}_a)$ of the continuous reaction rate so that

$$\tau(\mathbf{c}_a) = \tau^+(\mathbf{c}_a) - \tau^-(\mathbf{c}_a).$$

In order to ensure the positivity of c_M, the reaction rate r is defined by multiplying the negative part of τ by a control function.

$$r(\mathbf{c}) = \tau^+(\mathbf{c}_a) - H(c_M)\tau^-(\mathbf{c}_a), \tag{1.7}$$

where H is the Heaviside function defined in $[0, +\infty[$ by

$$\begin{cases} \text{if } x > 0 \text{ then } H(x) = 1, \\ \text{if } x = 0 \text{ then } H(x) = 0. \end{cases}$$

Clearly, the choice (1.7) satisfies (1.6). It can be noted that the Heaviside function H thus the reaction rate r can be defined for negative values: $H(x) = 0$ if $x < 0$. The kinetic model (1.2) is then a system of differential equations, where the reaction rate is a piecewise discontinuous function. It has a unique weak solution [12].

1.3.2 Set-valued model

The set-valued model does not specify the value of r with total dissolution. The idea is to limit the backward reaction rate β by a set-valued function:

$$r = \alpha a_1^{s_1} a_2^{s_2} - w_M \beta,$$

where $w_M \in W(c_M)$ with

$$\begin{aligned} W(u) &= \{1\} && \text{if } u > 0, \\ &= [0,1] && \text{if } u = 0, \\ &= \{0\} && \text{if } u < 0. \end{aligned} \tag{1.8}$$

Thus, $r = \tau$ as long as $c_M > 0$ and $\tau \leqslant r \leqslant \alpha a_1^{s_1} a_2^{s_2}$. When $c_M = 0$, the value of w_M can be chosen such that $r = 0$ [2]. This model can be extended to a set of aqueous and mineral reactions, provided that a mineral reaction contains exactly one mineral which participates only in this reaction of precipitation and dissolution [46]. An aqueous reaction R_j is written

$$\sum_{i \in \mathcal{A}_j^-} |s_{ij}| A_i \rightleftarrows \sum_{i \in \mathcal{A}_j^+} |s_{ij}| A_i,$$

whereas a mineral reaction R_j involving the mineral M_k is written

$$\sum_{i \in \mathcal{A}_j^-} |s_{ij}| A_i \rightleftarrows \sum_{i \in \mathcal{A}_j^+} |s_{ij}| A_i + M_k.$$

Let R_j be a reaction with mineral M_k, with τ_j defined by (1.3). Then the set-valued model for reaction rate r_j is written

$$r_j = \alpha_j \prod_{i \in \mathcal{A}_j^-} a_i(\mathbf{c}_a)^{|s_{ij}|} - w_k \beta_j \prod_{i \in \mathcal{A}_j^+} a_i(\mathbf{c}_a)^{|s_{ij}|},$$

where $w_k \in W(c_k)$ with the set-valued function W defined by (1.8).

Under some assumptions, the reactive transport model coupling advection-diffusion with the geochemical system has a unique weak solution [46]. The proof uses a regularized chemical system where the set-valued function W is replaced by a regularized Heaviside function.

1.3.3 Complementarity model

Here, the value of r when $c_M = 0$ is not given, since the function r is defined by the following complementarity problem:

$$c_M(r - \tau) = 0,\ c_M \geqslant 0,\ r \geqslant \tau,$$

which has a unique weak solution [46].

1.3.4 Reference example

In order to illustrate the appearance and disappearance of a pure phase, we study a very simple model with one reaction involving one aqueous species and one mineral: $A \rightleftharpoons M$.

System (1.2) is then written, with $\mathbf{c} = (c_A, c_M)$:

$$\begin{cases} \dfrac{\mathrm{d}c_A(t)}{\mathrm{d}t} = -r(c(t)), \\ \dfrac{\mathrm{d}c_M(t)}{\mathrm{d}t} = +r(c(t)), \\ c_A(t) \geqslant 0, c_M(t) \geqslant 0, \\ c_A(0) = c_{A,0}, c_M(0) = c_{M,0}. \end{cases} \tag{1.9}$$

The conservative variable is $(c_A + c_M)$, since $Q = (1, 1)^T$. Here, we choose a simple activity model with $a(c_A) = c_A$, so that the mass action law (1.3) becomes $\tau(c_A) = \alpha c_A - \beta$, which we rewrite (with $\alpha = 1$ and $\beta = K$)

$$\tau(c_A) = c_A - K,$$

where K is a strictly positive constant.

The continuous problem (1.4) with the reaction rate $\tau(c_A)$ has a unique solution:

$$\begin{cases} c_A^*(t) = K + \tau_0 \exp(-t), t \geqslant 0, \\ c_M^*(t) = c_{M,0} + \tau_0(1 - \exp(-t)), t \geqslant 0. \end{cases} \tag{1.10}$$

where $\tau_0 = c_{A,0} - K$. Thus $\tau(t) = \tau_0 \exp(-t)$ and $c_A^*(t) > 0, t > 0$.

If $\tau_0 = 0$, then $\tau(t) = 0$ and $r(t) = 0$, thus the system is at equilibrium at any time. From now on, we assume that $\tau_0 \neq 0$.

1.3.4.1 Solution of the discontinuous model

The discontinuous reaction rate (1.7) becomes here:

$$r(c_A, c_M) = \tau^+(c_A) - H(c_M)\tau^-(c_A). \tag{1.11}$$

The discontinuous model (1.9) with reaction rate (1.11) has a unique weak solution, satisfying definition 1.2.1. In what follows, we compute this solution, in the three situations depicted in Figure 1.1.

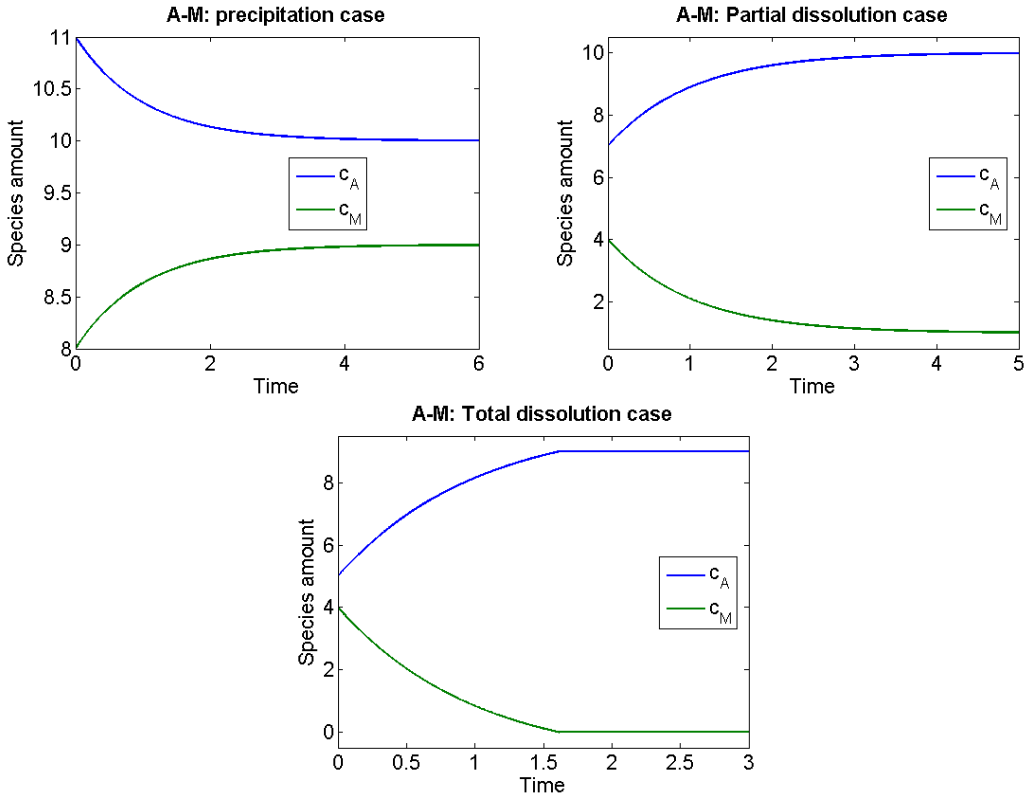


Figure 1.1 – Quantities of solute c_A and mineral c_M for reaction $A \rightleftharpoons M$ with $K = 10$. Top left: precipitation with $c_{A,0} = 11, c_{M,0} = 8$. Top right: partial dissolution with $c_{A,0} = 7, c_{M,0} = 4$. Bottom: full dissolution with $c_{A,0} = 5, c_{M,0} = 4$.

1.3.4.2 Precipitation

If $\tau_0 > 0$, then $\forall t > 0, c_A^*(t) > 0, c_M^*(t) > 0, \tau(t) > 0$. The solute precipitates but does not disappear and the quantity of mineral increases. Here, c^* is the weak solution of system (1.9) and formula (1.11) gives $r = \tau$.

1.3.4.3 Partial dissolution

If $-c_{M,0} \leqslant \tau_0 < 0$, then $\forall t > 0, c_A^*(t) > 0, c_M^*(t) > 0$. The mineral dissolves but does not disappear and the quantity of solute increases. Here too, c^* is the weak solution of system (1.9) and $r = \tau$.

1.3.4.4 Full dissolution

If $\tau_0 < -c_{M,0}$, then $c_M^*(t_M) = 0$ where $t_M \geqslant 0$ is given by

$$t_M = \log\left(\frac{\tau_0}{\tau_0 + c_{M,0}}\right) \tag{1.12}$$

The mineral is fully dissolved. Let (c_A, c_M) be defined by:

$$\begin{cases} c_A(t) = c_A^*(t), t \leqslant t_M, \\ c_M(t) = c_M^*(t), t \leqslant t_M, \\ c_A(t) = c_{A,0} + c_{M,0}, t \geqslant t_M, \\ c_M(t) = 0, t \geqslant t_M, \end{cases}$$

It is the weak solution of (1.9) with $r = H(c_M)\tau$.

1.4 A differential inclusion model for general geochemical systems

In this section, we consider the general geochemical system described in section 1.2. We develop an original new model based on a differential inclusion [39]. Another model defined in [34] is based on a projected dynamical system. In our model, the reaction rate r_j is equal to τ_j if the quantities of all reactant minerals in the reaction are strictly positive and it must be controlled if the quantity of one reactant mineral is zero. However, these controls may depend on the vector $\mathbf{c}_p$ of all mineral quantities. As in the discontinuous model, we limit the positive and negative parts of τ_j; nevertheless, these controls belong to set-valued functions as in the set-valued model.

Our objective is to define reaction rates $\mathbf{r}$ in order to ensure a non-negative solution. We propose a model where reaction rates r_j are governed by maximal reaction rates τ_j in the following way:

$$\forall j = 1, \ldots, N_r, \ |r_j| \leqslant |\tau_j| \text{ and } r_j\tau_j \geqslant 0. \tag{1.13}$$

For any $j \in \{1, \ldots, N_r\}$, we introduce set-valued functions U_j and V_j of variables $\mathbf{c}_p \in (\mathbb{R}^+)^{N_p}$. Then we introduce two control functions $u_j, v_j \in L^\infty(0,T)$ in order to define the reaction rates.

Definition 1.4.1. *The set-valued functions $U_j, j = 1, \ldots, N_r$ are defined in $[0, +\infty)^{N_p}$ by:*

$$\begin{aligned} U_j(\boldsymbol{c}_p) &= \{1\} \text{ if } \forall k \in \mathcal{M}_j^-, c_k > 0, \text{ or } \mathcal{M}_j^- = \varnothing, \\ &= [0,1] \text{ else.} \end{aligned}$$

The set-valued functions $V_j, j = 1, \ldots, N_r$ are defined in $[0, +\infty)^{N_p}$ by:

$$\begin{aligned} V_j(\boldsymbol{c}_p) &= \{1\} \text{ if } \forall k \in \mathcal{M}_j^+, c_k > 0, \text{ or } \mathcal{M}_j^+ = \varnothing, \\ &= [0,1] \text{ else.} \end{aligned}$$

For general geochemical systems, we can now define our model of reaction rates.

Definition 1.4.2. *Let $T > 0$ and $\boldsymbol{c} \in (H^1(0,T))^{N_a+N_p}$ such that $\boldsymbol{c}(t) \geqslant 0$, $0 \leqslant t \leqslant T$. Reaction rates $\boldsymbol{r}$ are defined by*

$$r_j = u_j\tau_j^+(\boldsymbol{c}) - v_j\tau_j^-(\boldsymbol{c}), \ j = 1, \ldots, N_r, \tag{1.14}$$

where $u_j, \ v_j \in L^\infty(0,T)$ and $u_j(t) \in U_j(\boldsymbol{c}_p(t))$, $v_j(t) \in V_j(\boldsymbol{c}_p(t))$, $0 < t < T$.

Clearly, the vector of reaction rates $\mathbf{r}$ satisfies (1.13). The reaction rate r_j, $j = 1, \ldots, N_r$ is equal to the maximal value τ_j if all the dissolving minerals are present. If, for example, $\tau_j > 0$ and a reactant mineral is fully dissolved, then $0 \leqslant r_j \leqslant \tau_j$. The reaction is not stopped but it may continue at a limited rate.

The problem to solve is then a differential inclusion system [39]:
find $\mathbf{c} \in (H^1(0,T))^{N_a+N_p}$, $\mathbf{u} \in (L^\infty(0,T))^{N_r}, \mathbf{v} \in (L^\infty(0,T))^{N_r}$ such that

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = S\boldsymbol{r}(t, \mathbf{c}(t)),\ t \in (0,T), \\ \mathbf{c}(t) \geqslant 0, \\ \mathbf{c}(0) = \mathbf{c}_0, \\ u_j(t) \in U_j(\mathbf{c}_p(t)), v_j(t) \in V_j(\mathbf{c}_p(t)),\ t \in (0,T), \\ r_j = u_j \tau_j^+(\mathbf{c}_a) - v_j \tau_j^-(\mathbf{c}_a),\ j = 1, \ldots, N_r, \end{cases} \tag{1.15}$$

where $\mathbf{c}_0 \geqslant 0$ is a given initial condition, τ is defined by (1.3), U_j and V_j are introduced in Definition 1.4.1. In order to analyze and to solve this system, we introduce a regularization method, as in [46]. A regularized model is controlled by a parameter ϵ, with a continuous reaction rate. We get a system of stiff ODEs and prove that it has a unique non negative solution. Then we prove that the differential inclusion (1.15) has a weak solution, limit of a convergent sequence of regularized functions. Moreover, the regularized problem can be solved by a classical ODE solver and provides a numerical method to approximate a solution of the discontinuous system.

1.4.1 Regularized model

In order to analyze the inclusion differential system, we introduce a regularized problem, as in [46]. Let us first define a regularized Heaviside function H_ϵ:

$$\begin{cases} \text{if } x \geqslant \epsilon \text{ then } H_\epsilon(x) = 1, \\ \text{if } 0 \leqslant x \leqslant \epsilon \text{ then } H_\epsilon(x) = \gamma(x/\epsilon), \\ \text{if } x = 0 \text{ then } H_\epsilon(x) = 0, \end{cases} \tag{1.16}$$

where γ is an increasing Lipschitz continuous function defined in the interval $[0,1]$ such that $\gamma(0) = 0$ and $\gamma(1) = 1$, so that the function H_ϵ is Lipschitz continuous and bounded. For example, we can choose $\gamma(x) = x$.

We can now define regularized reaction rates. If only one mineral M_k participates as a reactant in reaction R_j then $u_{j,\epsilon}(\mathbf{c}) = H_\epsilon(c_k)$ and $v_{j,\epsilon}(\mathbf{c}) = 1$. If it is a product, then $u_{j,\epsilon}(\mathbf{c}) = 1$ and $v_{j,\epsilon}(\mathbf{c}) = H_\epsilon(c_k)$. But if two minerals participate as reactants in the same reaction R_j, we use a blending function to define the limiter, which takes values in the interval $[0,1]$. This blending of the reactant minerals must be equal to 0 as soon as one of them disappears and be equal to 1 as long as all of them are present in sufficient quantity. Similar properties apply to product minerals. Examples of such blending are the minimum and the product functions. In what follows, we choose the min function.

Definition 1.4.3. *For general geochemical systems, the regularized reaction rates are defined by*

$$r_{j,\epsilon}(\boldsymbol{c}) = u_{j,\epsilon}(\boldsymbol{c})\tau_j^+(\boldsymbol{c}) - v_{j,\epsilon}(\boldsymbol{c})\tau_j^-(\boldsymbol{c}), \tag{1.17}$$

where $u_{j,\epsilon}$ and $v_{j,\epsilon}$ are defined by

$$\begin{cases} \text{if } \mathcal{M}_j^- = \varnothing, \text{ then } u_{j,\epsilon}(\boldsymbol{c}) = 1, \text{ else } u_{j,\epsilon}(\boldsymbol{c}) = \min\limits_{k \in \mathcal{M}_j^-}(H_\epsilon(c_k)), \\ \text{if } \mathcal{M}_j^+ = \varnothing, \text{ then } v_{j,\epsilon}(\boldsymbol{c}) = 1, \text{ else } v_{j,\epsilon}(\boldsymbol{c}) = \min\limits_{k \in \mathcal{M}_j^+}(H_\epsilon(c_k)). \end{cases} \tag{1.18}$$

It can be noted that these regularized reaction rates satisfy the requirements (1.13). We study the regularized system of kinetic reactions:

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{c}(t)}{\mathrm{d}t} = S\boldsymbol{r_\epsilon}(\mathbf{c}(t)), t \geqslant 0, \\ \mathbf{c}(0) = \mathbf{c}_0. \end{cases} \tag{1.19}$$

It is easy to show that the regularized system (1.19) has a unique solution.

Lemma 1.4.1. *The function r_ϵ is locally Lipschitz continuous.*

Proof. The functions $u_{j,\epsilon}$ and $v_{j,\epsilon}$ are bounded and Lipschitz continuous. The function τ is C^1 and locally bounded. Therefore the function r_ϵ is locally Lipschitz continuous. □

Proposition 1.4.1. *$\exists T > 0$ such that system (1.19) with non-negative initial condition $\boldsymbol{c}_0$, has a unique maximal non-negative solution $\boldsymbol{c}_\epsilon$ in $[0, T[$.*

Proof. We apply the theorem of Cauchy-Lipschitz. □

Now, in order to prove the non negativity of the solution, we start by showing that the right-hand side is essentially non negative.

Lemma 1.4.2. *The function f defined by $f(\boldsymbol{c}) = Sr_\epsilon(\boldsymbol{c})$ is locally Lipschitz continuous and essentially non negative.*

Proof. By using Lemma 1.4.1 and linearity, f is locally Lipschitz continuous. Let $\mathbf{c}$ a vector of species quantities such that $\mathbf{c} \geqslant 0$.

First, let $A_k, 1 \leqslant k \leqslant N_a$ be an aqueous species such that $c_k = 0$ for some $k \in \{1, \dots N_a\}$. We proved in Proposition 1.2.2 that $s_{kj}\tau_j \geqslant 0$, so using 1.13, $s_{kj}r_{j,\epsilon} \geqslant 0$ and $(Sr_\epsilon)_k \geqslant 0$.

Second, let $M_k, N_a + 1 \leqslant k \leqslant N_a + N_p$ be a mineral such that $c_k = 0$. If M_k is a reactant of reaction R_j, then $s_{kj} < 0$; since $c_k = 0$ we get $u_{j,\epsilon} = 0$ and $r_{j,\epsilon} \leqslant 0$. On the contrary, if M_k is a product of reaction R_j, then $s_{kj} > 0$; since $c_k = 0$ we get $v_{j,\epsilon} = 0$ and $r_{j,\epsilon} \geqslant 0$. Finally $(Sr_{j,\epsilon})_k \geqslant 0$. □

Proposition 1.4.2. *If $\boldsymbol{c}_0 \geqslant 0$ then the solution of system (1.19) satisfies $\boldsymbol{c}_\epsilon \geqslant 0$ in the maximal interval.*

Proof. As in Proposition 1.2.3, it is a direct consequence of Lemmas 1.4.1 and 1.4.2. □

1.4.2 Convergence of the regularized solutions

In this section, we prove convergence of the regularized solutions, towards a non-negative and bounded weak solution of the differential inclusion problem. From now on, we make the following assumption:

Assumption 1.4.1. *We assume the existence of a vector* $\boldsymbol{q}$ *such that* $\boldsymbol{q} > 0$ *and* $S^T\boldsymbol{q} = 0$.

This assumption implies that the regularized solution is bounded and can be defined in $[0, +\infty[$:

Theorem 1.4.1. *Under Assumption 1.4.1, system (1.19) has a unique solution* $\boldsymbol{c}_\epsilon$ *in* $[0, +\infty[$ *which satisfies*

$$0 \leqslant \boldsymbol{c}_\epsilon \leqslant C,$$

where C *is defined in Lemma 1.2.1.*

Proof. Let $\mathbf{c}_\epsilon$ be the maximal solution of (1.19). Since $S^T\mathbf{q} = 0$, it satisfies $\mathbf{q}^T\mathbf{c}_\epsilon = \mathbf{q}^T\mathbf{c}_0$. Thanks to Theorem 1.4.2, it also satisfies $\mathbf{c}_\epsilon \geqslant 0$. Thus, using Lemma 1.2.1, it also satisfies $\mathbf{c}_\epsilon \leqslant C$. So, the maximal interval of the non-negative solution is $[0, +\infty[$. □

It also implies that the sequence $\mathbf{c}_\epsilon$ converges:

Proposition 1.4.3. *Let* $T > 0$. *Under Assumption 1.4.1, the sequence* $\boldsymbol{c}_\epsilon$ *converges strongly in* $C^0(0, T, \mathbb{R}^{(N_a+N_p)})$. *Any limit* $\boldsymbol{c}$ *of a subsequence satisfies* $0 \leqslant \boldsymbol{c} \leqslant C$.

Also for any $j = 1, \ldots, N_r$, *subsequences of the regularized controls* $u_{j,\epsilon}(\boldsymbol{c}_\epsilon)$ *and* $v_{j,\epsilon}(\boldsymbol{c}_\epsilon)$ *converge weakly-* in* $L^\infty(0, T, \mathbb{R}^{N_r})$ *towards respectively* u_j *and* v_j *which satisfy* $0 \leqslant u_j \leqslant 1$ *and* $0 \leqslant v_j \leqslant 1$. *Moreover, for any* $j = 1, \ldots, N_r$ *and any* $0 < t < T$, *limits* $\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}$ *defined above satisfy* $u_j(t) \in U_j(\boldsymbol{c}_p(t))$ *and* $v_j(t) \in V_j(\boldsymbol{c}_p(t))$.

Proof. Let $\mathbf{c}_\epsilon$ be the solution of (1.19) in $[0, +\infty[$. Since $0 \leqslant \mathbf{c}_\epsilon \leqslant C$, the norms $\|\mathbf{c}_\epsilon\|_{L^2(0,T)}$ and $\|Sr_\epsilon(\mathbf{c}_\epsilon)\|_{L^2(0,T)}$ are bounded, thus $\|\mathbf{c}_\epsilon\|_{H^1(0,T)}$ is also bounded.

By compact injection of $H^1(0, T)$ in $C^0(0, T)$, a subsequence of $\mathbf{c}_\epsilon$ converges in $C^0(0, T)$ and the limit $\mathbf{c}$ is bounded below by 0 and above by C.

Since $0 \leqslant u_{j,\epsilon}(\mathbf{c}_\epsilon) \leqslant 1$, a subsequence of $u_{j,\epsilon}(\mathbf{c}_\epsilon)$ converges weakly-* in $L^\infty(0, T)$ and the limit u_j is bounded below by 0 and above by 1. The result is similar for v_j. Thus $\forall j = 1, \ldots, N_r$, the control satisfies $u_j \in L^\infty(0, T)$ and $0 \leqslant u_j(t) \leqslant 1$, $0 < t < T$.

Now, we want to show that $u_j(t) \in U_j(\mathbf{c}_p(t))$. Let $1 \leqslant j \leqslant N_r$ and $0 < t^* < T$ such that $\mathcal{M}_j^- \neq \varnothing$ and $\forall k \in \mathcal{M}_j^-$, $c_k(t^*) > 0$. Since $\mathbf{c}$ is continuous, there exists an interval I containing t^* such that $\forall t \in I, \forall k \in \mathcal{M}_j^-$, $c_k(t) > 0$. Since a subsequence of $\mathbf{c}_\epsilon$ converges towards $\mathbf{c}$ in $C^0(0, T)$, for ϵ sufficiently small, the same property is satisfied: $\forall t \in I, \forall k \in \mathcal{M}_j^-$, $c_{k,\epsilon}(t) > \delta_k = \frac{c_k(t)}{2} > 0$. Consequently, $\forall t \in I, u_{j,\epsilon}(\mathbf{c}_\epsilon)(t) = 1$, so that $\lim_{\epsilon \to 0} u_{j,\epsilon}(\mathbf{c}_\epsilon) = 1$ in $C^0(I)$, and $\forall t \in I, u_j(t) = 1$. This concludes the proof. □

Now, we can prove the existence of a weak solution of system (1.15):

Theorem 1.4.2. *We make Assumption 1.4.1. Let* $T > 0$ *and* $\boldsymbol{c}$ *a limit of* $\boldsymbol{c}_\epsilon$ *in* $C^0(0, T, \mathbb{R}^{(N_a+N_p)})$. *Let* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *weakly-* limits of a subsequence of* $\boldsymbol{u}_\epsilon(\boldsymbol{c}_\epsilon)$ *and* $\boldsymbol{v}_\epsilon(\boldsymbol{c}_\epsilon)$ *in* $L^\infty(0, T, \mathbb{R}^{N_r})$. *Then* $\boldsymbol{c}, \boldsymbol{u}, \boldsymbol{v}$ *are a weak solution of system (1.15).*

Proof. First, the limits u_j and v_j satisfy the requirements of definition 1.4.2. Also, $\mathbf{c}(0) = \mathbf{c}_0$ and $\mathbf{c} \geqslant 0$. Since $\mathbf{c}_\epsilon$ weakly converges towards $\mathbf{c}$ in $H^1(0,T)$, $\forall \phi \in (\mathcal{C}_c^\infty(0,T))$,

$$\lim_{\epsilon \to 0} \int_0^T \big(\frac{\mathrm{d}\mathbf{c}_\epsilon(s)}{\mathrm{d}t} - \frac{\mathrm{d}\mathbf{c}(s)}{\mathrm{d}t}\big)\phi(s)\ \mathrm{d}s = 0. \tag{1.20}$$

Thanks to Lemma 1.4.3, $u_{j,\epsilon}, u_j \in L^\infty(0,T)$ and $\|u_{j,\epsilon}\|_\infty \leqslant 1$. Also, since τ_j and $\mathbf{c}_\epsilon, \mathbf{c}$ are continuous, $\tau_j^+(\mathbf{c}), \tau_j^+(\mathbf{c}_\epsilon) \in L^1(0,T)$.

Let us decompose $r_{j,\epsilon} - r_j$ in the following way:

$$\begin{array}{rcl} r_{j,\epsilon} - r_j & = & u_{j,\epsilon}(\mathbf{c}_\epsilon)(\tau_j^+(\mathbf{c}_\epsilon) - \tau_j^+(\mathbf{c})), \\ & + & (u_{j,\epsilon}(\mathbf{c}_\epsilon) - u_j)\tau_j^+(\mathbf{c}), \\ & - & v_{j,\epsilon}(\mathbf{c}_\epsilon)(\tau_j^-(\mathbf{c}_\epsilon) - \tau_j^-(\mathbf{c})), \\ & - & (v_{j,\epsilon}(\mathbf{c}_\epsilon) - v_j)\tau_j^-(\mathbf{c}). \end{array}$$

Then $\forall \phi \in \mathcal{C}_c^\infty(0,T)$,

$$\begin{array}{rcl} |\int_0^T (r_{j,\epsilon}(\mathbf{c}_\epsilon(s)) - r_j(s,\mathbf{c}(s)))\phi(s)\ \mathrm{d}s| & \leqslant & \|\tau_j^+(\mathbf{c}_\epsilon) - \tau_j^+(\mathbf{c})\|_{L^1(0,T)}\|\phi\|_{L^\infty(0,T)} \\ & + & |\int_0^T (u_{j,\epsilon}(\mathbf{c}_\epsilon(s)) - u_j(s))\tau_j^+(\mathbf{c}(s))\phi(s)\ \mathrm{d}s| \\ & + & \|\tau_j^-(\mathbf{c}_\epsilon) - \tau_j^-(\mathbf{c}))\|_{L^1(0,T)}\|\phi\|_{L^\infty(0,T)} \\ & + & |\int_0^T (v_{j,\epsilon}(\mathbf{c}_\epsilon(s)) - v_j(s))\tau_j^-(\mathbf{c}(s))\phi(s)\ \mathrm{d}s|. \end{array}$$

Since $\mathbf{c}_\epsilon$ converges towards $\mathbf{c}$ in $C^0(0,T)$ and τ_j is continuous, we get $\lim_{\epsilon \to 0} \|\tau_j^+(\mathbf{c}_\epsilon) - \tau_j^+(\mathbf{c})\|_{L^1(0,T)} = 0$. Since a subsequence of $u_{j,\epsilon}(\mathbf{c}_\epsilon)$ weakly-* converges to u_j in $L^\infty(0,T)$ and $\tau_j^+(\mathbf{c})\phi \in L^1(0,T)$, we also obtain $\lim_{\epsilon \to 0} \int_0^T (u_{j,\epsilon}(\mathbf{c}_\epsilon) - u_j)\tau_j^+(\mathbf{c})\phi\ \mathrm{d}s = 0$. Similar results apply to v_j and τ_j^-.

So, $\forall \phi \in \mathcal{C}_c^\infty(0,T)$,

$$\lim_{\epsilon \to 0} \int_0^T (Sr_\epsilon(\mathbf{c}_\epsilon(s)) - Sr(s,\mathbf{c}(s)))\phi(s)\ \mathrm{d}s = 0. \tag{1.21}$$

Using both weak convergence results (1.20) and (1.21), we conclude that $\mathbf{c}, \mathbf{u}, \mathbf{v}$ are a weak solution of system (1.15).

□

1.5 Test case with three minerals

The problem we consider here is a set of two reactions with three minerals: $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$. This system is similar to the two reactions between calcite, calcium oxide and calcium hydroxide. The two continuous reaction rates defined by the mass action laws (1.3) are constant:

$$\tau = (\tau_1, \tau_2).$$

System (1.15) becomes

$$\begin{cases} \dfrac{dc_1}{dt} = -r_1, \\ \dfrac{dc_2}{dt} = r_1 - r_2, \\ \dfrac{dc_3}{dt} = r_2, \\ c(0) = c_0, \\ c \geqslant 0, \end{cases} \tag{1.22}$$

with $r_1 = u_1\tau_1^+ - v_1\tau_1^-$ and $r_2 = u_2\tau_2^+ - v_2\tau_2^-$.

The kernel of S^T is of dimension 1 and $q = (1,1,1)^T$ is a basis vector satisfying $q > 0$. The conservative variable is $c_1 + c_2 + c_3$ so that it is possible to consider only two equations out of the three in (1.22).

The unique solution c^* of system (1.4) is a linear function of time:

$$\begin{cases} c_1^*(t) = c_{1,0} - \tau_1 t, \\ c_2^*(t) = c_{2,0} + (\tau_1 - \tau_2)t, \\ c_3^*(t) = c_{3,0} + \tau_2 t. \end{cases}$$

It is easy to compute the first times $t_i^* \geqslant 0$ when the quantities c_i^* are zero (and then negative):

$$\begin{cases} \text{If } \tau_1 > 0 \text{ then } t_1^* = \dfrac{c_{1,0}}{\tau_1}, \\ \text{If } \tau_2 - \tau_1 > 0 \text{ then } t_2^* = \dfrac{c_{2,0}}{\tau_2 - \tau_1}, \\ \text{If } \tau_2 < 0 \text{ then } t_3^* = -\dfrac{c_{3,0}}{\tau_2}. \end{cases} \tag{1.23}$$

Let $c_{\epsilon,i}^*$ be the solution of the regularized system (1.19), where the function γ is identity. We analyze its convergence in different cases. We will show stronger results than Theorem 1.4.2. Indeed, c_ϵ converges towards c and for any $t > 0$, the regularized controls $u_\epsilon(c_\epsilon(t))$ and $v_\epsilon(c_\epsilon(t))$ converge towards discontinuous functions $u(c(t))$ and $v(c(t))$. Moreover, c is the unique solution of a discontinuous system.

1.5.1 Dissolution of M_1 and M_3

Here, we assume that $\tau_1 > 0$ and $\tau_2 < 0$, that $c_{1,0} > 0$ and $c_{3,0} > 0$. The pure phase M_1 dissolves and produces M_2 while M_3 dissolves and produces M_2. The regularized reaction rates are given by

$$\begin{cases} r_{1,\epsilon} = H_\epsilon(c_1)\tau_1, \\ r_{2,\epsilon} = H_\epsilon(c_3)\tau_2. \end{cases} \tag{1.24}$$

The regularized solutions are easy to compute in that case, since the first and third equations are independent.

Taking ϵ sufficiently small, let $t_{1,\epsilon} = \dfrac{c_{1,0} - \epsilon}{\tau_1}$ so that $0 < t_{1,\epsilon} < t_1^*$ and $\lim t_{1,\epsilon} = t_1^*$. The solution $c_{1,\epsilon}$ is then

$$\begin{cases} c_{1,\epsilon}(t) &= c_{1,0} - \tau_1 t = c_1^*(t), & t \leqslant t_{1,\epsilon}, \\ c_{1,\epsilon}(t) &= \epsilon \exp(-\dfrac{\tau_1}{\epsilon}(t - t_{1,\epsilon})), & t \geqslant t_{1,\epsilon}. \end{cases}$$

It is easy to verify that $\forall t \geqslant 0$, $\lim_{\epsilon \to 0} c_{1,\epsilon}(t) = c_1(t)$ where c_1 is defined by

$$\begin{cases} c_1(t) &= c_{1,0} - \tau_1 t = c_1^*(t), & t \leqslant t_1^*, \\ c_1(t) &= 0, & t \geqslant t_1^*. \end{cases}$$

Also, the control converges: $\forall t \geqslant 0, t \neq t_1^* \lim_{\epsilon \to 0} H_\epsilon(c_{1,\epsilon}(t)) = H(c_1(t))$.

Clearly, results are the same for the variable c_3 and c_2 is deduced from the conservative quantity. Thus c is solution of (1.22) where

$$r_1 = H(c_1)\tau_1, \; r_2 = H(c_3)\tau_2.$$

Figure 1.2 illustrates this case. Mineral M_2 precipitates at rate $(\tau_1 - \tau_2)$ until M_3 is fully dissolved, then reaction R_2 stops and M_2 precipitates at rate τ_1 until M_1 is fully dissolved and reaction R_1 stops.

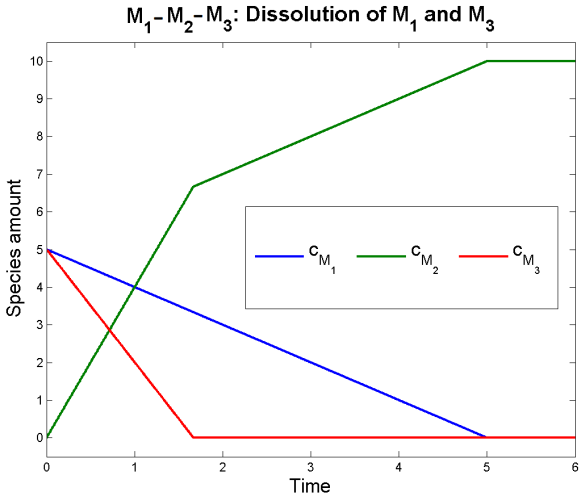


Figure 1.2 – Reactions $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$ with maximal rates $\tau_1 = 1, \tau_2 = -3$ and initial conditions $c_{1,0} = 5, c_{2,0} = 0, c_{3,0} = 5$.

1.5.2 Dissolution of M_2

Here, we assume that $\tau_1 < 0$ and $\tau_2 > 0$, that $c_{2,0} > 0$. The pure phase M_2 dissolves and produces M_1 and M_3. The regularized reactions rates are given by

$$\begin{cases} r_{1,\epsilon} = H_\epsilon(c_2)\tau_1, \\ r_{2,\epsilon} = H_\epsilon(c_2)\tau_2. \end{cases} \tag{1.25}$$

In that case, the second equation is independent and similar to the previous case. In the same way as $t_{1,\epsilon}$, we define $t_{2,\epsilon} = \dfrac{c_{2,0} - \epsilon}{\tau_2 - \tau_1}$ so that $0 < t_{2,\epsilon} < t_2^*$ and $t_{2,\epsilon}$ converges to t_2^*. Thus the quantity $c_{2,\epsilon}$ converges for all $t \geqslant 0$ towards c_2 defined as c_1 above (with $(\tau_2 - \tau_1)$ instead of τ_1), and the control $H_\epsilon(c_{2,\epsilon}(t))$ converges towards $H(c_2(t))$ for $t \neq t_2^*$.

It is slightly more complicated to compute the solution $c_{1,\epsilon}$:

$$\left\{\begin{array}{rcll} c_{1,\epsilon}(t) & = & c_{1,0}-\tau_1 t = c_1^*(t), & t \leqslant t_{2,\epsilon}, \\ c_{1,\epsilon}(t) & = & -\epsilon\dfrac{\tau_1}{\tau_2-\tau_1}(1-\exp(-\dfrac{\tau_2-\tau_1}{\epsilon}(t-t_{2,\epsilon}))) + c_1^*(t_{2,\epsilon}), & t \geqslant t_{2,\epsilon}. \end{array}\right.$$

Therefore, for all $t \geqslant 0$, the regularized solution converges towards c_1 defined by

$$\left\{\begin{array}{rcll} c_1(t) & = & c_{1,0}-\tau_1 t = c_1^*(t), & t \leqslant t_2^*, \\ c_1(t) & = & c_1^*(t_2), & t \geqslant t_2^*. \end{array}\right.$$

The third quantity c_3 is obtained thanks to the conservation law. The limit $\mathbf{c}$ is solution of (1.22) where reaction rates are the discontinuous functions

$$r_1 = H(c_2)\tau_1,\ r_2 = H(c_2)\tau_2.$$

As illustrated in Figure 1.3, minerals M_1 and M_3 precipitate until M_2 is fully dissolved and both reactions R_1 and R_2 stop.

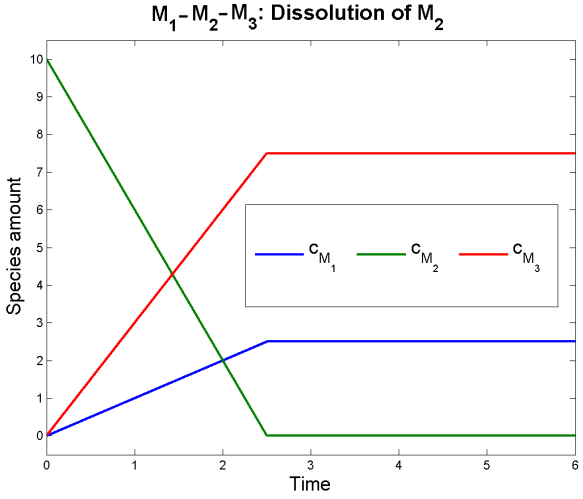


Figure 1.3 – Reactions $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$ with maximal rates $\tau_1 = -1, \tau_2 = 3$ and initial conditions $c_{1,0} = c_{3,0} = 0, c_{2,0} = 10$.

1.5.3 Dissolution of M_1 then M_2

Here, we assume that $\tau_2 > 0$ and $\tau_1 > 0$, the mineral M_1 dissolves and produces M_2 while M_2 dissolves and produces M_3. As previously, the regularized solution $c_{1,\epsilon}$ converges towards c_1 defined above. We also assume that $t_{2,\epsilon} < 0$ or $t_{1,\epsilon} \leqslant t_{2,\epsilon}$.

With these assumptions, we get, as long as $c_{2,\epsilon}(t) \geqslant \epsilon$ $H_\epsilon(c_{2,\epsilon}) = 1$ and

$$\left\{\begin{array}{rcll} c_{2,\epsilon}(t) & = & c_{2,0}-(\tau_2-\tau_1)t = c_2^*(t), & t \leqslant t_{1,\epsilon}, \\ c_{2,\epsilon}(t) & = & \epsilon(1-\exp(-\dfrac{\tau_1}{\epsilon}(t-t_{1,\epsilon}))) - \tau_2(t-t_{1,\epsilon}) + c_2^*(t_{1,\epsilon}), & t \geqslant t_{1,\epsilon}. \end{array}\right.$$

On the other hand, using that $c_{2,\epsilon}(t_{1,\epsilon}) \geqslant \epsilon$,

$$\lim_{t\to+\infty} \epsilon(1-\exp(-\frac{\tau_1}{\epsilon}(t-t_{1,\epsilon}))) - \tau_2(t-t_{1,\epsilon}) + c_2^*(t_{1,\epsilon}) = -\infty$$

and $t \mapsto \epsilon(1-\exp(-\frac{\tau_1}{\epsilon}(t-t_{1,\epsilon}))) - \tau_2(t-t_{1,\epsilon}) + c_2^*(t_{1,\epsilon})$ is decreasing for all $t \geqslant t_{1,\epsilon}$, there exists a unique $\hat{t}_{2,\epsilon} \geqslant t_{1,\epsilon}$ such that $c_{2,\epsilon}(\hat{t}_{2,\epsilon}) = \epsilon$. Then, for $t \geqslant \hat{t}_{2,\epsilon}$ the equation becomes

$$\frac{dc_2}{dt} = \frac{c_1}{\epsilon}\tau_1 - \frac{c_2}{\epsilon}\tau_2,\ t \geqslant \hat{t}_{2,\epsilon}$$

with $c_2(\hat{t}_{2,\epsilon}) = \epsilon$. The solution is

$$c_{2,\epsilon}(t) = \epsilon(a\exp(-\frac{\tau_1}{\epsilon}(t-t_{1,\epsilon})) + (1-ab_\epsilon)\exp(-\frac{\tau_2}{\epsilon}(t-\hat{t}_{2,\epsilon})))$$

where $a = \dfrac{\tau_1}{\tau_2-\tau_1}$ and $b_\epsilon = \exp(-\frac{\tau_1}{\epsilon}(\hat{t}_{2,\epsilon}-t_{1,\epsilon}))$.

We get

$$\begin{cases} c_{2,\epsilon}(t) &=& c_{2,0} - (\tau_2-\tau_1)t = c_2^*(t), & t \leqslant t_{1,\epsilon}, \\ c_{2,\epsilon}(t) &=& \epsilon(1-\exp(-\frac{\tau_1}{\epsilon}(t-t_{1,\epsilon}))) - \tau_2(t-t_{1,\epsilon}) + c_2^*(t_{1,\epsilon}), & t_{1,\epsilon} \leqslant t \leqslant \hat{t}_{2,\epsilon}, \\ c_{2,\epsilon}(t) &=& \epsilon(a\exp(-\frac{\tau_1}{\epsilon}(t-t_{1,\epsilon})) + (1-ab_\epsilon)\exp(-\frac{\tau_2}{\epsilon}(t-\hat{t}_{2,\epsilon}))) & t \geqslant \hat{t}_{2,\epsilon} \end{cases}$$

Let us now analyze the convergence of $\hat{t}_{2,\epsilon}$ and $c_{2,\epsilon}(t)$.

We recall below the definition of $\hat{t}_{2,\epsilon}$

$$c_{2,\epsilon}(\hat{t}_{2,\epsilon}) = \epsilon,$$

direct computations yield the following relation for all ϵ

$$-\epsilon\exp(-\frac{\tau_1}{\epsilon}(\hat{t}_{2,\epsilon}-t_{1,\epsilon}))) - \tau_2(\hat{t}_{2,\epsilon}-t_{1,\epsilon}) + c_2^*(t_{1,\epsilon}) = 0,$$

since $\hat{t}_{2,\epsilon} > t_{1,\epsilon}$ for all ϵ we have that $\exp(-\frac{\tau_1}{\epsilon}(\hat{t}_{2,\epsilon}-t_{1,\epsilon}))$ is uniformly bounded for all ϵ, hence the first term vanishes when $\epsilon \to 0$. This ensures the existence of $\lim\limits_{\epsilon\to 0}\hat{t}_{2,\epsilon}$ since $t_{1,\epsilon}$ and $c_2^*(t_{1,\epsilon})$ admit a limit. We set then $\lim\limits_{\epsilon\to 0}\hat{t}_{2,\epsilon} = \hat{t}_2$. Moreover, in order to compute the limit c_2 we show that $\hat{t}_{2,\epsilon} \leqslant \hat{t}_2$. Indeed, the function

$$g(t,\epsilon) := \epsilon(1-\exp(-\frac{\tau_1}{\epsilon}(t-t_{1,\epsilon}))) - \tau_2(t-t_{1,\epsilon}) + c_2^*(t_{1,\epsilon})$$

is decreasing in ϵ for all $t \geqslant t_1^*$. This ensures that for all $\epsilon > \epsilon'$ and $t \geqslant t_1^*$, $g(t,\epsilon) < g(t,\epsilon')$. Moreover, note that $g(\cdot,\epsilon) = c_{2,\epsilon}$ on $[t_{1,\epsilon}, \hat{t}_{2,\epsilon}]$ (which decreases toward $-\infty$ when t goes to $+\infty$), then necessarily $\hat{t}_{2,\epsilon} < \hat{t}_{2,\epsilon'}$ for all $\epsilon > \epsilon'$. Passing through the limit $\epsilon' \to 0$ we get $\hat{t}_{2,\epsilon} \leqslant \hat{t}_2$ for all $\epsilon > 0$.

We get $\lim\limits_{\epsilon\to 0}\hat{t}_{2,\epsilon} = \hat{t}_2$ which satisfies $\hat{t}_{2,\epsilon} \leqslant \hat{t}_2$ and $\tau_2(\hat{t}_2 - t_1^*) = c_2^*(t_1^*)$. Then $\forall t \geqslant 0, \lim\limits_{\epsilon\to 0} c_{2,\epsilon}(t) = c_2(t)$ given by

$$\begin{cases} c_2(t) &=& c_2^*(t), & t \geqslant t_1^*, \\ c_2(t) &=& -\tau_2(t-t_1^*) + c_2^*(t_1^*), & t_1^* \leqslant t \leqslant \hat{t}_2, \\ c_2(t) &=& 0, & t \geqslant \hat{t}_2. \end{cases}$$

Moreover, $\forall t \geqslant 0, t \neq t_1^*, t \neq \hat{t}_2$, $\lim_{\epsilon \to 0} H_\epsilon(c_{2,\epsilon}(t)) = H(c_2(t))$.

Finally, $\mathbf{c}$ is solution of the discontinuous system (1.22) with

$$r_1 = H(c_1)\tau_1,\ r_2 = H(c_2)\tau_2.$$

In Figure 1.4 where $t_2^* < 0$, mineral M_1 dissolves at rate τ_1 and mineral M_2 dissolves at rate $\tau_2 - \tau_1$ then reaction R_1 stops and M_2 dissolves at rate τ_2 until reaction R_2 stops.

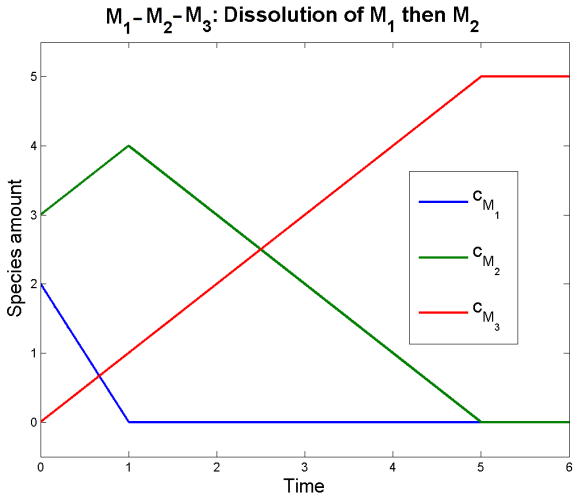


Figure 1.4 – Reactions $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$ with maximal rates $\tau_1 = 2, \tau_2 = 1$ and initial conditions $c_{1,0} = 2, c_{2,0} = 3, c_{3,0} = 0$.

1.5.4 Dissolution of M_2 then M_1

Up to now, the regularized Heaviside functions converged towards the Heaviside function and reactions stopped when the reactant mineral was fully dissolved. Here, we will see that reactions can continue although M_2 is fully dissolved.

We assume that $\tau_2 > \tau_1 > 0$ and $t_2^* < t_1^*$ so that, for ϵ sufficiently small, $0 < t_{2,\epsilon} \leqslant t_{1,\epsilon}$.

The second regularized equation is

$$\frac{dc_2}{dt} = H_\epsilon(c_1)\tau_1 - H_\epsilon(c_2)\tau_2.$$

It is still possible to compute the regularized solution $c_{2,\epsilon}$, although we do not provide all the details here:

$$\begin{cases} c_{2,\epsilon}(t) &= c_{2,0} - (\tau_2 - \tau_1)t = c_2^*(t), & t \leqslant t_{2,\epsilon}, \\ c_{2,\epsilon}(t) &= \dfrac{\epsilon}{\tau_2}(\tau_1 + (\tau_2 - \tau_1)\exp(-\dfrac{\tau_2}{\epsilon}(t - t_{2,\epsilon}))), & t_{2,\epsilon} \leqslant t \leqslant t_{1,\epsilon}, \\ c_{2,\epsilon}(t) &= \epsilon(b_{1,\epsilon}\exp(-\dfrac{\tau_2}{\epsilon}(t - t_{2,\epsilon})) + b_{2,\epsilon}\exp(-\dfrac{\tau_2}{\epsilon}(t - t_{1,\epsilon})) & \\ &+ b_{3,\epsilon}\exp(-\dfrac{\tau_1}{\epsilon}(t - t_{2,\epsilon})) + b_{4,\epsilon}\exp(-\dfrac{\tau_1}{\epsilon}(t - t_{1,\epsilon}))), & t \geqslant t_{1,\epsilon}, \end{cases}$$

where $b_{i,\epsilon}$ are bounded epsilon-dependant constants. Thus the regularized solution $c_{2,\epsilon}$ converges, for all $t \geqslant 0$, towards c_2 defined by

$$\begin{cases} c_2(t) &= c_{2,0} - (\tau_2 - \tau_1)t = c_2^*(t), & t < t_2^*, \\ c_{2,\epsilon}(t) &= 0, & t \geqslant t_2^*. \end{cases}$$

Also, the regularized control $H_\epsilon(c_{2,\epsilon})$ has a special limit, for any $t \geqslant 0$, $t \neq t_1^*, t \neq t_2^*$:

$$\lim_{\epsilon \to 0} H_\epsilon(c_{2,\epsilon}(t)) = H(c_2(t)) + H(c_1(t))\delta(c_2(t))\frac{\tau_1}{\tau_2},$$

where $\delta(x)$ is the Dirac function at point x. The limit c of the regularized solution satisfies the discontinuous differential equations (1.22) with

$$r_1 = H(c_1)\tau_1,\ r_2 = H(c_2)\tau_2 + H(c_1)\delta(c_2)\tau_1.$$

Mineral M_3 precipitates at rate τ_2 until M_2 is fully dissolved. Then, reaction R_2 does not stop but continues at rate τ_1, so that M_2 remains fully dissolved, until M_1 is fully dissolved. In some sense, the two reactions are replaced by the unique reaction $M_1 \to M_3$ at rate τ_1. This behaviour is shown in Figure 1.5.

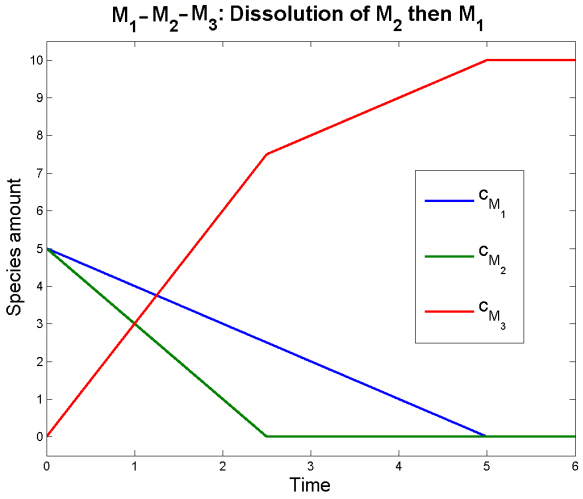


Figure 1.5 – Reactions $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$ with maximal rates $\tau_1 = 1, \tau_2 = 3$ and initial conditions $c_{1,0} = 5, c_{2,0} = 5, c_{3,0} = 0$.

1.5.5 Summary of test case $M_1 - M_2 - M_3$

The other cases can be easily deduced from the four cases considered above. Finally, the regularized solution converges towards c which is solution of a discontinuous differential system. Moreover, the regularized controls converge, except at a finite number of discontinuity points, to the discontinuous ones.

Theorem 1.5.1. *The limit c of the regularized solution c_ϵ is solution of system (1.22) where the reaction rates are discontinuous functions defined by:*

$$\begin{cases} r_1 = H(c_1)\tau_1^+ - H(c_2)\tau_1^- - \delta(c_2)H(c_3)\tau_2^-, \\ r_2 = H(c_2)\tau_2^+ + \delta(c_2)H(c_1)\tau_1^+ - H(c_3)\tau_2^-. \end{cases} \tag{1.26}$$

Moreover, the regularized limiters $u_\epsilon(c_\epsilon)$ and $v_\epsilon(c_\epsilon)$ strongly converge, up to a subsequence, in $L^p(0,T)$, $p \in [1, \infty[$.

Proof. Existence is proved above by construction. We proved in all cases that $u_\epsilon(c_\epsilon)$ converges almost everywhere to u in $[0,T]$. Thanks to the dominated convergence theorem applied to $|u_\epsilon(c_\epsilon(t)) - u(t)|^p$, we get strong convergence in L^p. □

We observe, in all cases of this example, that when a mineral disappears at a given time, then it remains fully dissolved afterwards. In other words, if $c_i(t^*) = 0$ then $c_i(t) = 0, t \geqslant t^*$ so that $\dfrac{dc_i}{dt} = 0, t > t^*$. Finally, the reaction rates satisfy $(Sr)_i = 0, t > t^*$. For example, when c_2 becomes fully dissolved, then $r_2 = r_1$.

1.6 Numerical experiments

In this section we describe and discuss the results of numerical experiments done with synthetic geochemical systems. We developed a Matlab code for solving the regularized ODE (1.19) and we study numerical convergence by varying the value of the parameter.

We define six test cases. The first one is the reference example $A - M$ defined in Section 1.3.4. The second one is $M_1 - M_2 - M_3$ analyzed in Section 1.5. Then we replace M_1 and M_3 by aqueous species A_1 and A_3 in the third test case. The fourth one (respectively the fifth one) corresponds to the previous second test case (respectively previous third test case) for which we add one mineral species M_4 and one reaction $M_2 \rightleftharpoons M_4$ (respectively one aqueous species A_4 and one reaction $M_2 \rightleftharpoons A_4$). With this construction, both reactions consuming M_2 in both, fourth and fifth, cases are in competition. Finally, the last test case contains seven minerals and five reactions such that one of them contains two minerals in the reactive side. In this configuration, we show numerically the non-uniqueness of the limiter $\boldsymbol{u}$ according the the choice of the regularized limiter $\boldsymbol{u_\epsilon}$.

1.6.1 Numerical context

1.6.1.1 ODE solver

In our experiments, we used the linear function defined in (1.16). We also implemented an exponential smoothing and a polynomial one with similar results. The value of ϵ is a parameter that varies to analyze the numerical convergence.

For cases with more than one mineral per reaction we introduced a blending function (1.18) in the regularized model. In the first experiments, there is no need of blending. In the last one, we compare the min and the prod functions, in order to observe the effect of blending. The blending product is defined as follows

$$\begin{cases} \text{if } \mathcal{M}_j^- = \varnothing, \text{ then } u_{j,\epsilon}(\mathbf{c}) = 1, \text{ else } u_{j,\epsilon}(\mathbf{c}) = \displaystyle\prod_{k \in \mathcal{M}_j^-} (H_\epsilon(c_k)), \\ \text{if } \mathcal{M}_j^+ = \varnothing, \text{ then } v_{j,\epsilon}(\mathbf{c}) = 1, \text{ else } v_{j,\epsilon}(\mathbf{c}) = \displaystyle\prod_{k \in \mathcal{M}_j^+} (H_\epsilon(c_k)). \end{cases} \tag{1.27}$$

We used the Matlab solver ode15s which implements an implicit scheme with variable order in order to solve stiff differential equations (see [67] and Matlab documentation for ode15s). An adaptive time step uses tolerance options *RelTolSolver* and *AbsTolSolver*, both taken as the default value 10^{-6}. The Jacobian matrix of the nonlinear right-hand

side is computed by a finite difference method. The time interval $[0, T]$ and the initial condition $\boldsymbol{c_0} \geqslant 0$ are parameters.

We define a set of discrete times in $[0, T]$: $t_k = k\ dt,\ k = 0, \cdots, N_f$ with $t_{N_f} = T$ and $dt = \dfrac{T}{N_f}$ where dt is an external time step. We get as output of ode15s a set of discrete values $\boldsymbol{c_\epsilon}(t_k)$, where $\boldsymbol{c_\epsilon}(t_k)$ is a vector of size $N_a + N_p$.

1.6.1.2 Physical model

The maximal reaction rates are implemented in a Matlab function, using data of the physical model, mainly stoichiometry and constants of equilibrium. In our tests, the activity of a mineral species is equal to 1 and the activity of an aqueous species is equal to the molar fraction:

$$a_i(\boldsymbol{c_a}) = \frac{c_i}{c_{\mathrm{H_2O}} + \sum_{i=1}^{N_a} c_{a,i}}, \tag{1.28}$$

where $c_{\mathrm{H_2O}}$ is the quantity of water. This definition assumes that the coefficient of activity is equal to 1. Moreover, we consider a dilute solution such that $c_{\mathrm{H_2O}}$ is considered constant.

1.6.1.3 Discrete error norms

In some tests, we know the limit $\boldsymbol{c}$ of the regularized model, so that $\boldsymbol{c}$ is an exact solution of the differential inclusion model. Therefore we can compute a discrete error for each species i:

$$e_i(t_k) = |c_i(t_k) - c_{\epsilon,i}(t_k)|,\ k = 0, \cdots, N_f,$$

and a global error:

$$e(t_k) = \sum_{i=1}^{N_a+N_p} e_i(t_k).$$

Then we compute the L^p-norm with $p = 1, 2, \infty$.

We also analyze the errors in the limiting discontinuous functions $u_{\epsilon,j}$ and $v_{\epsilon,j}$ for each reaction R_j. In some tests, we also know the limits u_j and v_j of $u_{\epsilon,j}$ and $v_{\epsilon,j}$. The number N_l of discontinuity points is finite and the corresponding times $t_k^*,\ k = 1, \cdots, N_l$, are known. As above, we compute discrete error for each reaction R_j and L^p-norm:

$$e_{u_j}(t_k) = |u_j(t_k) - u_{\epsilon,j}(t_k)|,\ k = 0, \cdots, N_f,$$

$$e_{v_j}(t_k) = |v_j(t_k) - v_{\epsilon,j}(t_k)|,\ k = 0, \cdots, N_f.$$

We also compute the discrete error at times t_k outside of $[t_k^* - \delta, t_k^* + \delta]$, $\forall k = 1, \cdots, N_l$, where $\delta > 0$. We call out-of-jumps norms the L^p-norms computed in this subset.

1.6.2 Reference example $A - M$, total dissolution

In this section we consider the reference example of Section 1.3.4:

$$A \rightleftharpoons M,$$

with A a solute in the water phase $w := \{H_2O, A\}$ and M a mineral. The stoichiometric matrix is $S = \begin{pmatrix} -1 & 1 \end{pmatrix}^T$.

We consider the total dissolution case, as in Section 1.3.4.4, with a different reaction rate. The initial state is $c_0 = (0.2\ 0.6)$ and $c_{H_2O} = 2$. In particular, the water is not involved in the reaction hence the quantity of water is constant (its value is 2). The activity of M is equal to 1 (as every mineral). The activity of A is defined by (1.28) and the reaction rate by (1.3), so that

$$a(c_A) := \frac{c_A}{c_A + c_{H_2O}} \text{ and } \tau(c_A) := \alpha a(c_A) - \beta$$

Table 1.1 contains a summary of physical parameters of this test case.

Reaction	$A \rightleftharpoons M$
S	$\begin{pmatrix} -1 \\ 1 \end{pmatrix}$
(α, β)	$(2, 2)$
c_{H_2O}	2
c_0	$(0.2\ 0.6)$

Table 1.1 – Test case $A - M$: physical parameters

Figure 1.6 shows the regularized solution and the reaction rate computed with $\epsilon = 0.001$.

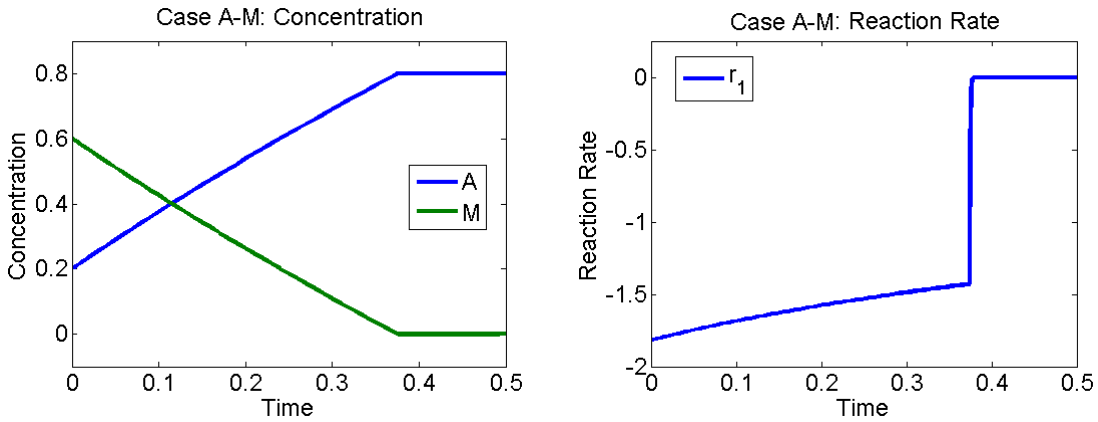


Figure 1.6 – Case $A - M$ with dissolution of M. Results with $\epsilon = 0.001$.

Physical data are such that the mineral is fully dissolved at time t^*. As expected from Section 1.3, for $t \geqslant t^*$ the quantity of mineral c_M remains equal to 0 and the reaction stops: $r = \tau$ until time t^* then $r = 0$.

Now we run the regularized model for a decreasing sequence of parameters ϵ and compute errors. We choose $T = 0.5$, $dt = \dfrac{T}{10^4} = 5 \times 10^{-5}$ and $\delta = 0.1$. Results are given in Figure 1.7.

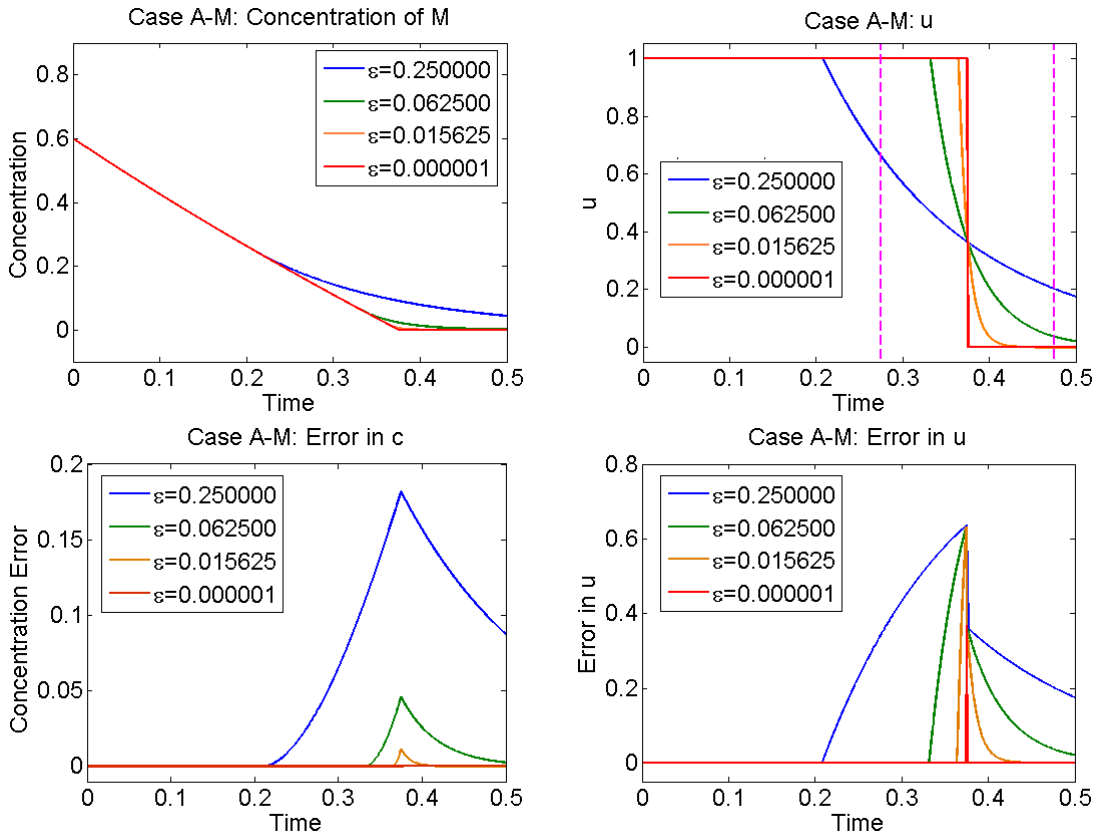


Figure 1.7 – Case $A - M$ with dissolution of M. Results and errors for different values of ϵ. Top: Quantity of mineral (left) and function u (right). Bottom: error e (left) and error e_u (right). The interval $[t^* - \delta, t^* + \delta]$ is represented in pink.

We can observe that errors are localized around the discontinuity. For the concentrations, the maximum value of the error and the length of the interval around t^* for which the error is not negligible decreases when ϵ tends towards 0. For the limiter u, the area decreases when ϵ tends towards 0 but the error remains large at the discontinuity point. Indeed, $\lim_{\epsilon \to 0} H_\epsilon(c_{M,\epsilon}(t)) = H(c_M(t))$ outside a neighbourhood of t^*.

In Figure 1.8, we show the convergence of the concentration in L^1 and L^∞, as proved in Lemma 1.4.3. The order of convergence is about 1.85 for the L^1-norm and 1 for the L^∞-norm. For the limiter, the numerical result shows in Figure 1.8 are more general in this case. In fact, there is convergence in L^1 and in L^∞ outside the discontinuity. We do not represent the values of L^1-norm without the discontinuity but the convergence rate is faster than L^1-norm with the discontinuity. The convergence rates of the error for the seven first values of ϵ are almost 0.93 for the L^1-norm, this value increases towards 1 if we consider the five last values of ϵ.

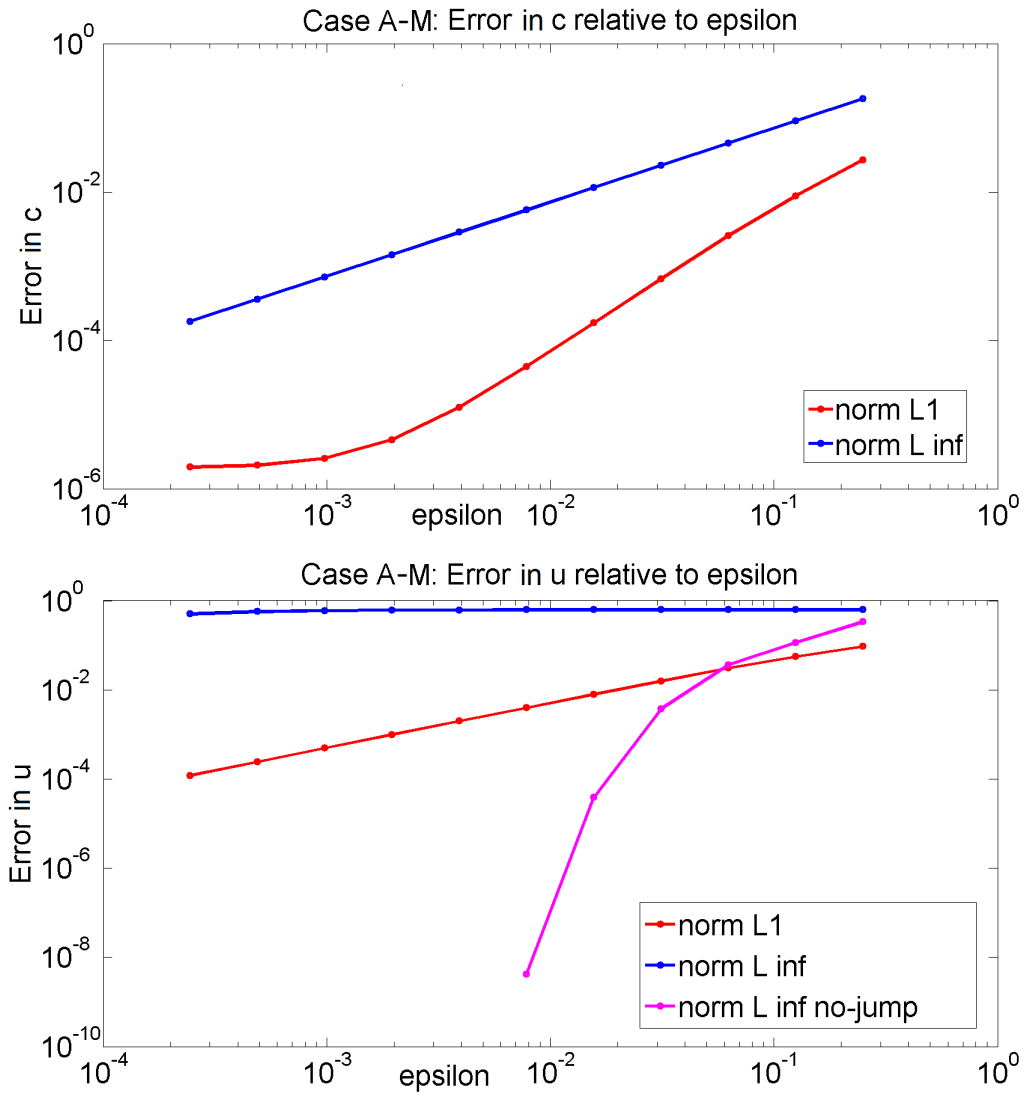


Figure 1.8 – Numerical convergence for the case $A - M$ with dissolution of M.

1.6.3 Test case $M_1 - M_2 - M_3$

The following example is described in Section 1.5,

$$M_1 \rightleftharpoons M_2 \text{ and } M_2 \rightleftharpoons M_3,$$

with M_1, M_2 and M_3 are three pure phase minerals. The stoichiometric matrix is $S = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}^T$. We consider the case described in Section 1.5.4, where a mineral, here M_2, precipitates in one reaction and is dissolved in a second one.

To compute the reaction rate we use $\alpha = (4\ 12)^T$ and $\beta = (3\ 8)^T$, both quantities are defined in (1.3). Then $\tau_1 = 1$ and $\tau_2 = 4$. In other words, this test case corresponds to the following choice for physical quantities: $K = (\frac{1}{4}\ \frac{1}{3}\ \frac{1}{2})^T$ and $k_f = k_b = (1\ 4)^T$. In this case, both reaction rates are positive. The initial state is $c_0 = (1\ 0.6\ 0.4)^T$, with all of this parameters, we are in the configuration of Section 1.5.4.

Table 1.2 contains a summary of physical parameters of this test case.

Reactions	$R_1 : M_1 \rightleftharpoons M_2$ $R_2 : M_2 \rightleftharpoons M_3$
S	$\begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$
(α_1, β_1) (α_2, β_2)	$(4, 3)$ $(12, 8)$
c_0	$(1, 0.6, 0.4)$

Table 1.2 – Test case $M_1 - M_2 - M_3$: physical parameters

First, we apply solver ode15s directly to the discontinuous system (1.26), with various input parameters. In any case, the solver fails, with a false result or no result at all. This example shows the need for a numerical regularization.

Figure 1.9 shows the regularized solution and both limiters computed with $\epsilon = 0.001$.

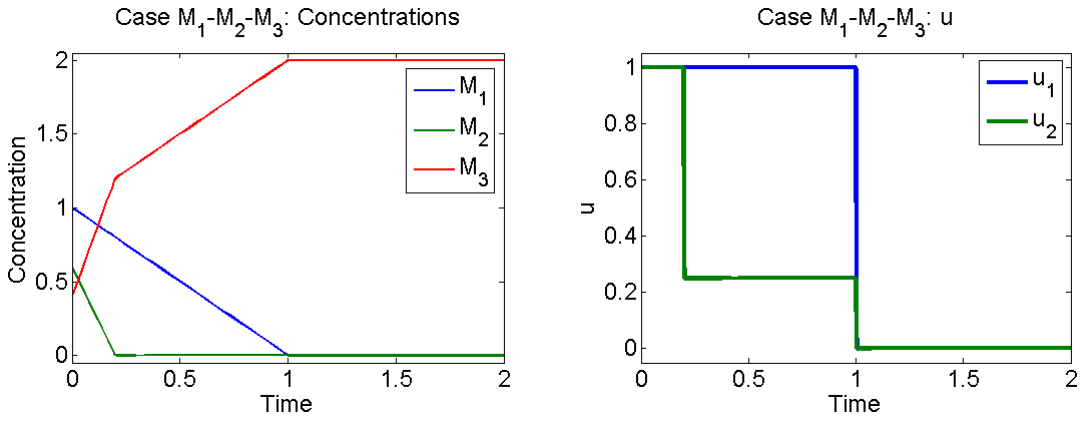


Figure 1.9 – Case $M_1 - M_2 - M_3$ with dissolution of M_2 then M_1. Results with $\epsilon = 0.001$.

We can see qualitatively that the solution obtained is close to the solution computed in Section 1.5.4. In order to justify this observation, we investigate the error of the regularized solution computed with our method, for a decreasing sequence of ϵ, according to the real one. Let us consider $T = 2$, $dt = \dfrac{2}{10^4} = 2 \times 10^{-4}$ and $\delta = 0.1$, defined in Section 1.6.

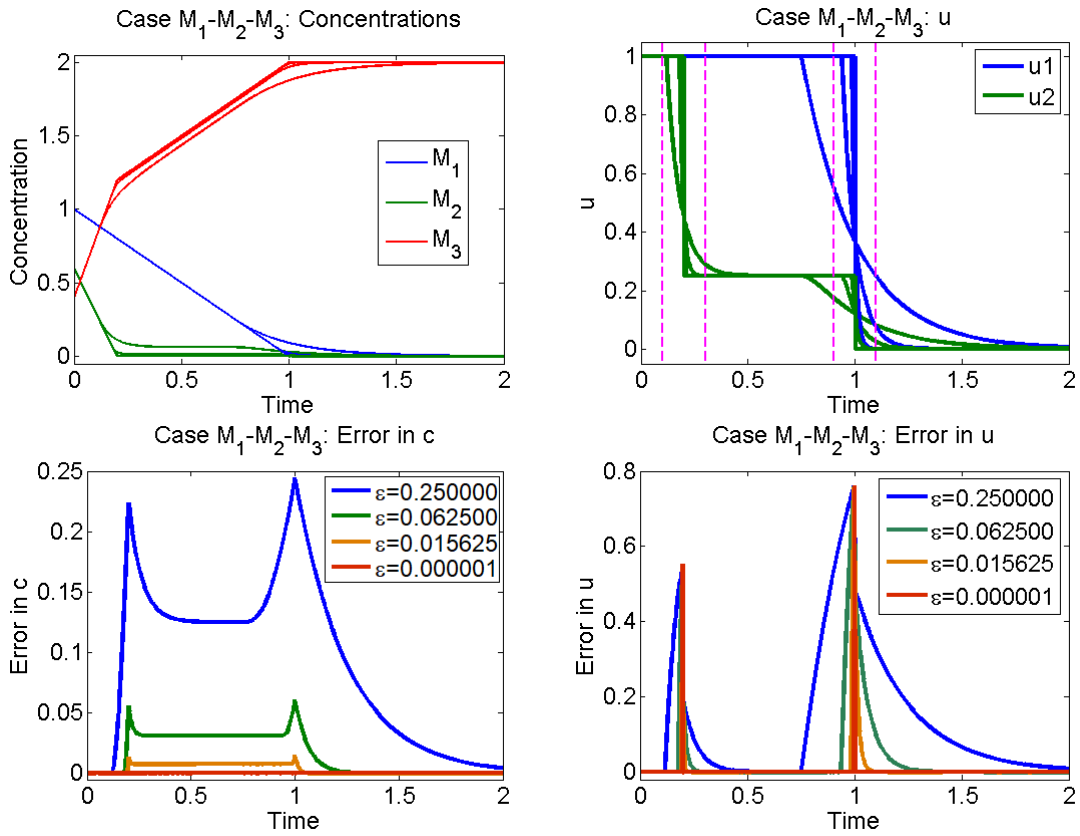


Figure 1.10 – Case $M_1 - M_2 - M_3$ with dissolution of M_2 then M_1. Results and errors for different values of ϵ. Top: concentrations (left) and function u (right). Bottom: error e (left) and error e_u (right). Both intervals $[t^* - \delta, t^* + \delta]$ are represented in pink.

In the graph representing the concentration, we observe a convergence when ϵ tends to 0. Moreover, the graph representing the error of the concentration shows that the error is maximal at the discontinuity times and the error converges to 0 when ϵ tends to 0.

Both graphs representing the limiter indicate that the error is localized around the discontinuities. The difference with the concentration is that the maximal values of the error do not seem to decrease.

Figure 1.11 represents errors in function of ϵ for $\boldsymbol{c}$ and $\boldsymbol{u}$. In fact, Theorem 1.4.2 ensures that c_ϵ converges strongly towards c in $\mathcal{C}^0$. Numerically, the rate of convergence is almost 1 for both norms.

For the limiter, Theorem 1.4.2 ensures only a weakly-* convergence in L^∞ but in this case we can prove the strong convergence in L^1. Numerically, the L^1-norm converges with a rate almost equal to 1 but the uniform convergence is not satisfied.

Without considering the discontinuities, the limiter converges uniformly. Numerically, we obtain a really fast convergence outside the jumps.

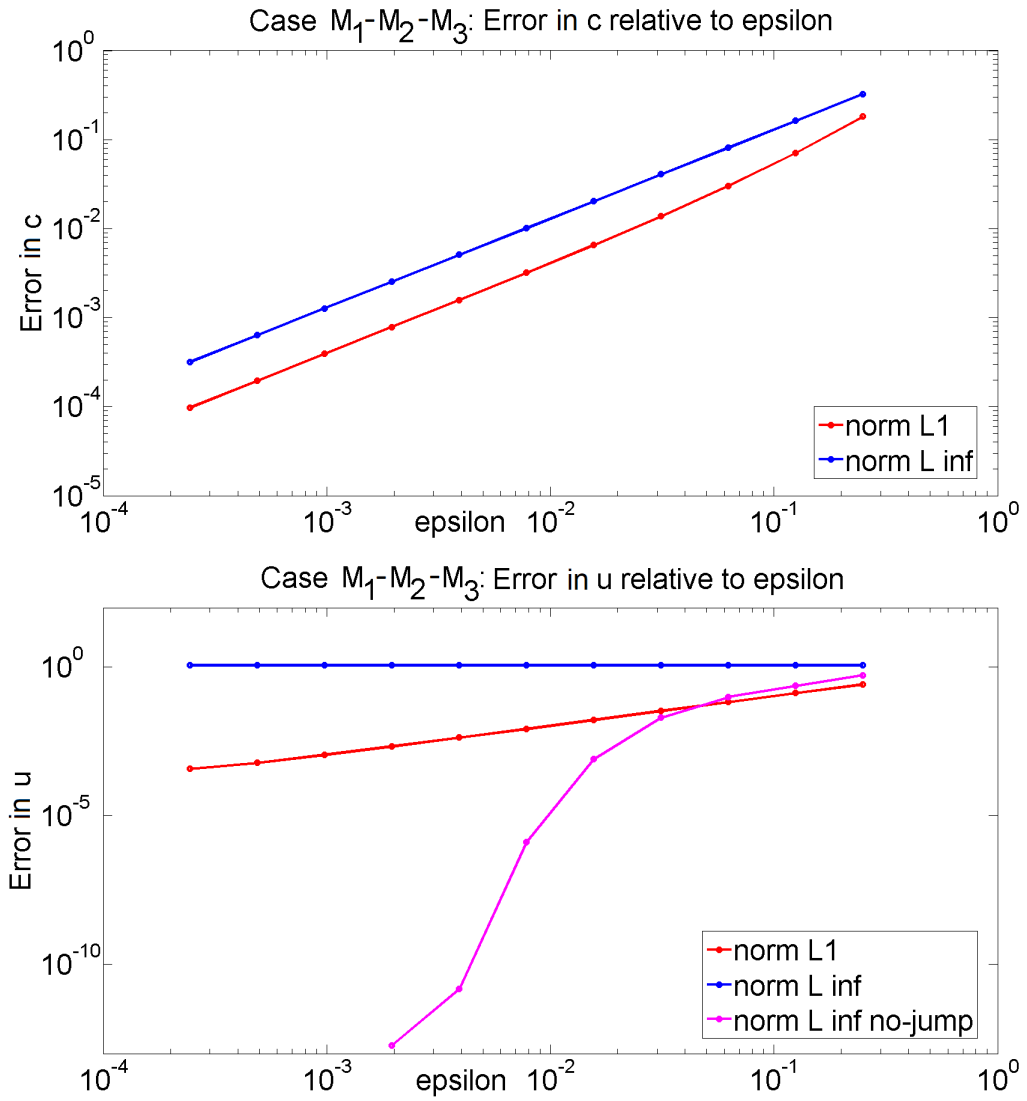


Figure 1.11 – Numerical convergence for the case $M_1 - M_2 - M_3$ in dissolution of M_2 then M_1.

1.6.4 Test case $A_1 - M_2 - A_3$

The following example has the same structure as the previous example. The difference here is the consideration of the aqueous species which results in the non-constant value of the reaction rates,

$$A_1 \rightleftharpoons M_2 \text{ and } M_2 \rightleftharpoons A_3,$$

where A_1 and A_3 are aqueous species (in this case the water phase is $w := \{H_2O, A_1, A_3\}$) and M_2 a pure phase mineral. The stoichiometric matrix is

$$S = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}^T,$$

with the first line corresponds to H_2O. In the sequel, we consider only the case where the reaction rates are positive and the M_2 species disappears before the first reaction reaches

equilibrium.

The reaction rates are computed using $\alpha = (30\ 4.5)^T$ and $\beta = (3\ 18.75)^T$, both quantities are defined in (1.3). In other word, this test case corresponds to the following choice for physical quantities: $K = (\frac{1}{1}\ \frac{1}{30}\ \frac{1}{3}\ \frac{2}{25})^T$ and $k_f = k_b = (1\ 1.5)^T$. The initial state is $c_0 = (2\ 5\ 0)^T$, the quantity of the solvent H_2O is more than 90 percent of the total initial quantity of the water phase ($c_{H_2O} = 20$).

The activity of water solute A_1 and A_3 is given by (1.28) and reaction rates by (1.3):

$$a_i(\boldsymbol{c_a}) = \frac{c_{A_i}}{c_{A_1} + c_{A_3} + c_{H_2O}},\ i = 1 \text{ or } 3,$$

$$\tau_1(\boldsymbol{c_a}) = \alpha_1 a_1(\boldsymbol{c_a}) - \beta_1 \text{ and } \tau_2(\boldsymbol{c_a}) = \alpha_2 - \beta_2 a_3(\boldsymbol{c_a})$$

Table 1.3 contains a summary of physical parameters of this test case.

Reactions	$R_1 : A_1 \rightleftharpoons M_2$ $R_2 : M_2 \rightleftharpoons A_3$
S	$\begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$
(α_1, β_1) (α_2, β_2)	$(30, 3)$ $(4.5, 18.75)$
c_0	$(2, 5, 0)$
c_{H_2O}	20

Table 1.3 – Test case $A_1 - M_2 - A_3$: physical parameters

Figure 1.12 shows the regularized solution and reaction rates computed with $\epsilon = 0.001$,

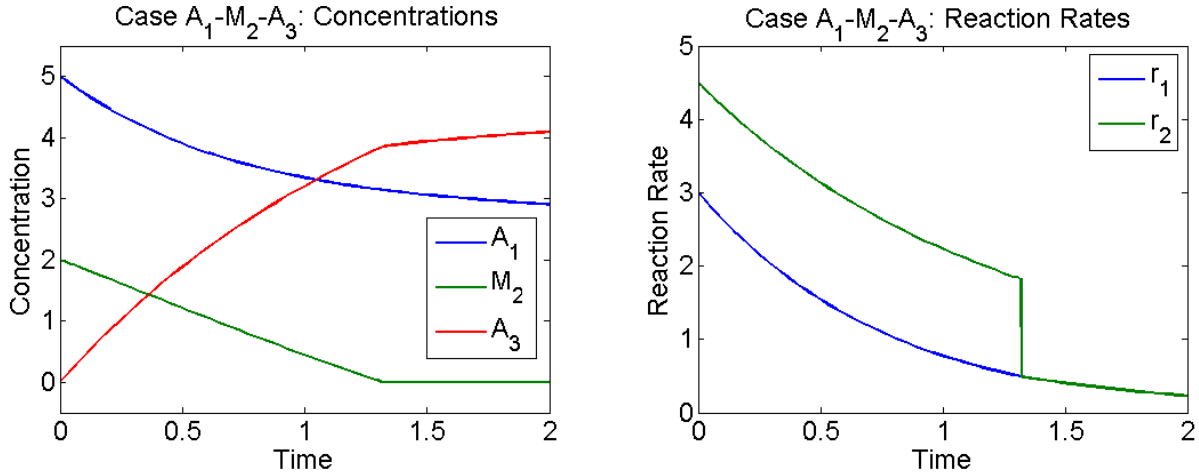


Figure 1.12 – Case $A_1 - M_2 - A_3$ with total dissolution of M_2. Results with $\epsilon = 0.001$.

We observe that, as in Section 1.6.3, when the central species M_2 disappears, the value of the limiter u_2 is such that $r_2 = r_1 = \tau_1$, i.e. $u_1 = 1$ and $u_2 = \dfrac{\tau_1}{\tau_2}$.

This example illustrates the behavior of limiters in the case where the reaction rates are not constant functions.

1.6.5 Test case with competition

The aim of the following examples are to observe reactions in competition. We want to specify the solution given by our model in the case where a reaction produces a mineral species and the two other reactions consume this one. The interesting case is when the species dissolved by the competition disappears. In fact, in the case $M_1 - M_2 - M_3$ we prove that reaction does not stop; the same result holds true in the case of competition. So, the main issue is to determine the new reaction rate value after the disappearance. We numerically prove that our model gives as value to the limiter the evenly distributed theoretical value (1.29).

We consider the following example, $M_1 \rightleftharpoons M_2$ and $M_2 \rightleftharpoons M_3$ and $M_2 \rightleftharpoons M_4$ with M_1, M_2, M_3 and M_4 four pure phase minerals. The stoichiometric matrix is $S = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}^T$. In the sequel we consider only the case where the reaction rates are positive and the M_2 species disappears first.

The reaction rates are computed using $\alpha = (4\ 1.5\ 6)^T$ and $\beta = (3\ 1\ 4)^T$, both quantities are defined in (1.3), so that $\tau_1 = 1$, $\tau_2 = 0.5$ and $\tau_3 = 2$. In other word, this test case corresponds to the following choice for physical quantities: $K = (\frac{1}{4}\ \frac{1}{3}\ \frac{1}{2}\ \frac{1}{2})^T$ and $k_f = k_b = (1\ 0.5\ 2)^T$. The initial state is $c_0 = (1\ 0.6\ 0.3\ 0.2)^T$.

Table 1.4 contains a summary of physical parameters of this test case.

Reactions	$R_1 : M_1 \rightleftharpoons M_2$ $R_2 : M_2 \rightleftharpoons M_3$ $R_3 : M_2 \rightleftharpoons M_4$
S	$\begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
(α_1, β_1) (α_2, β_2) (α_3, β_3)	$(4, 3)$ $(1.5, 1)$ $(6, 4)$
c_0	$(1, 0.6, 0.3, 0.2)$

Table 1.4 – Test case with competition: physical parameters

We recall that the limiters for the second and the third reaction are the same, as defined in (1.17).

Now, let us consider the linear regularization with $\epsilon = 0.001$. The results represented in Figure 1.13 allow us to extend those observed for the case $M_1 - M_2 - M_3$.

Here, mineral M_2 is fully dissolved at time $t_2^* = 0.4$, but reaction $\mathcal{R}_1$ still continues until time $t_1^* = 1$, when mineral M_1 is fully dissolved. We observe a numerical convergence

of the limiters. Between times t_2^* and t_1^*, we get $u_1 = 1$ and

$$u_2 = u_3 = u_1 \frac{\tau_1}{\tau_2 + \tau_3} = u_1 \frac{1}{0.5 + 2} = 0.4u_1, \text{ if } c_2 = 0. \tag{1.29}$$

After time t_1^*, we get $u_1 = u_2 = u_3 = 0$. In this example, reactions $\mathcal{R}_2$ and $\mathcal{R}_3$ are limited proportionally to the ratio of consumption and production.

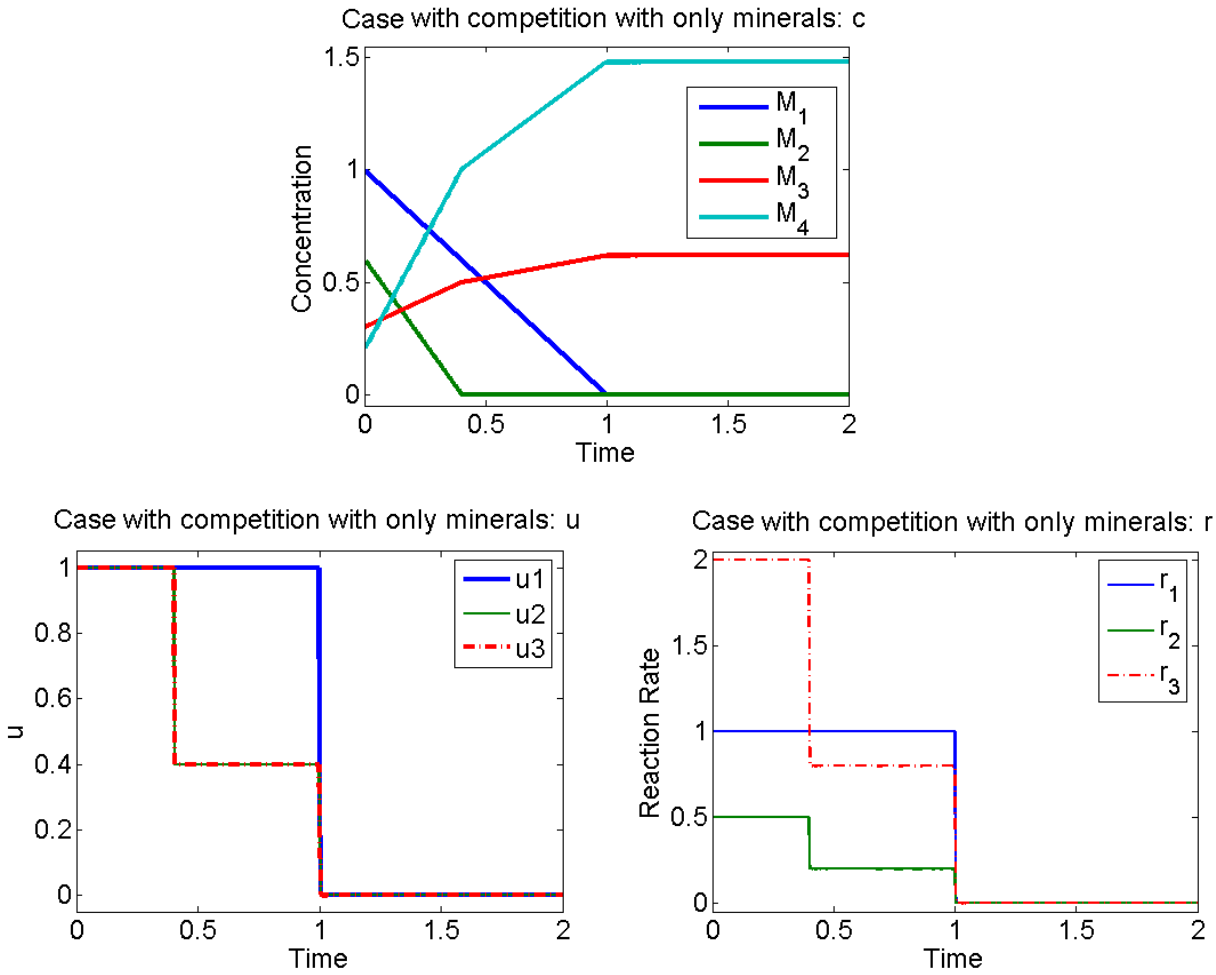


Figure 1.13 – Case with competition with only minerals. Results with $\epsilon = 0.001$.

Let us consider a similar test case including aqueous species as described below,

$$A_1 \rightleftharpoons M_2 \text{ and } M_2 \rightleftharpoons A_3 \text{ and } M_2 \rightleftharpoons A_4,$$

with M_2 a pure phase mineral and A_1, A_3 and A_4 are species in aqueous phase $w := \{H_2O, A_1, A_3, A_4\}$. The stoichiometric matrix is the same as before. The reaction rates are computed using $\alpha = (30\ 4.5\ 4.5)^T$ and $\beta = (3\ 18.75\ 45)^T$, both quantities are defined in (1.3). The activity of water solute A_1, A_3 and A_4 is given by (1.28) and reaction rates by (1.3):

$$a_i(\boldsymbol{c}_a) = \frac{c_{A_i}}{c_{A_1} + c_{A_3} + c_{A_4} + c_{H_2O}}, \ i = 1, \ 3 \text{ or } 4,$$

$$\tau_1(\boldsymbol{c_a}) = \alpha_1 a_1(\boldsymbol{c}_a) - \beta_1, \ \tau_2(\boldsymbol{c_a}) = \alpha_2 - \beta_2 a_3(\boldsymbol{c}_a) \text{ and } \tau_3(\boldsymbol{c_a}) = \alpha_3 - \beta_3 a_4(\boldsymbol{c}_a).$$

In this case, the three reaction rates are positive. The initial state is $c_0 = (5\ 2\ 0\ 0)^T$ and the intial quatity of H_2O is $c_{H_2O} = 20$.

Table 1.5 contains a summary of physical parameters of this test case.

Reactions	$R_1 : A_1 \rightleftharpoons M_2$ $R_2 : M_2 \rightleftharpoons A_3$ $R_3 : M_2 \rightleftharpoons A_4$
S	$\begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
(α_1, β_1) (α_2, β_2) (α_3, β_3)	$(30, 3)$ $(4.5, 18.75)$ $(4.5, 45)$
c_0	(5200)
c_{H_2O}	20

Table 1.5 – Test case with competition and aqueous species: physical parameters

Figure 1.14 shows the regularized solution and reaction rates computed with $\epsilon = 0.001$,

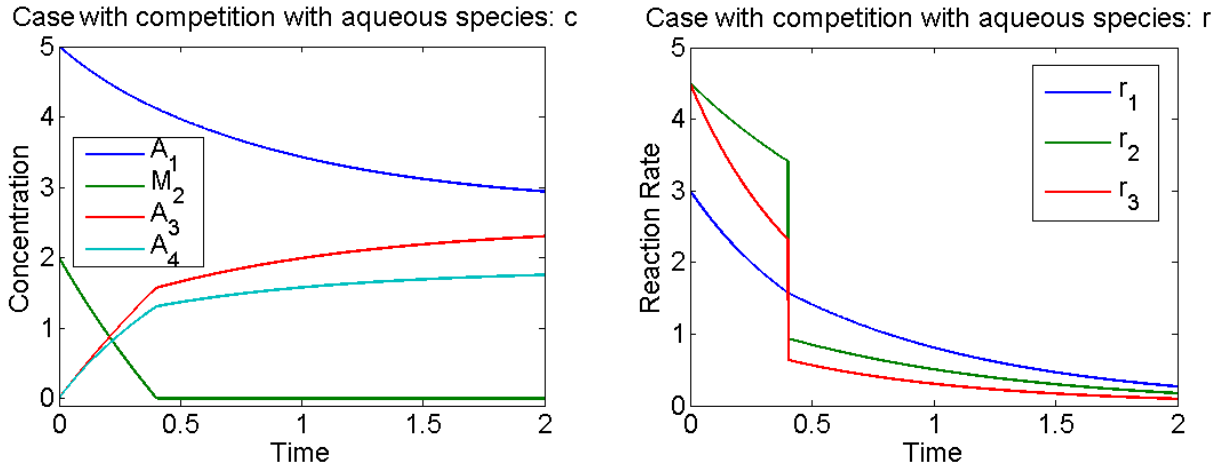


Figure 1.14 – Case Competition with aqueous species. Results with $\epsilon = 0.001$.

Numerically, we obtain a similar convergence of the limiters. Indeed, they converge towards u_2 and u_3 given by: $u_2 = u_3 = u_1 \frac{\tau_1}{\tau_2 + \tau_3}$, if $c_2 = 0$, with $r_i = u_i \tau_i$, for $i = 1, 2, 3$ and $u_1 = 1$.

1.6.6 Test case with blending (non-uniqueness)

The aim of this section is to show the non-uniqueness of the limiter. We consider the following example,

$$M_3 \rightleftharpoons M_1 \text{ and } M_4 \rightleftharpoons M_2 \text{ and } M_1 + M_2 \rightleftharpoons M_5 \text{ and } M_1 \rightleftharpoons M_6 \text{ and } M_2 \rightleftharpoons M_7,$$

with M_i, $i = 1, \cdots, 7$, are pure phase minerals. In the sequel we consider only the case where reaction rates are positive and M_1 and M_2 disappears first. The reaction rates are computed using $\alpha = (1.5\ 2.1\ 4\ 2\ 2)^T$ and $\beta = (1\ 1.4\ 3\ 1\ 1)^T$, both quantities are defined in (1.3),

$$\tau_1 = 0.5,\ \tau_2 = 0.7,\ \tau_3 = 1,\ \tau_4 = 1 \text{ and } \tau_5 = 1.$$

Table 1.6 contains a summary of physical parameters of this test case. Here, reaction $\mathcal{R}_3$ involves two minerals as reactants, so that the regularized limiter $u_{3,\epsilon}$ is defined with a blending function. We use here the min function as in (1.18) and the prod function as in (1.27).

Reactions	$R_1 : M_3 \rightleftharpoons M_1$ $R_2 : M_4 \rightleftharpoons M_2$ $R_3 : M_1 + M_2 \rightleftharpoons M_5$ $R_4 : M_1 \rightleftharpoons M_6$ $R_5 : M_2 \rightleftharpoons M_7$
S	$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$
(α_1, β_1)	$(1.5, 1)$
(α_2, β_2)	$(2.1, 1.4)$
(α_3, β_3)	$(4, 3)$
(α_4, β_4)	$(2, 1)$
(α_5, β_5)	$(2, 1)$
c_0	$(0.2, 0.1, 2, 2, 2, 2, 2)$

Table 1.6 – Test case with blending: physical parameters

Figure 1.15 shows numerical results obtained with eps=0.001 for these two blendings. We observe that limiters converge, but to different values.

Indeed, with min function, we get numerical convergence to

$$u_4 = 0.25,\ u_5 = 0.45,\ u_3 = \min(u_4, u_5) = 0.25. \tag{1.30}$$

With prod function, numerical limits of limiters are

$$u_4 = \frac{\sqrt{\frac{43}{2}} - 3}{5} \simeq 0.32736,\ u_5 = \frac{\sqrt{\frac{43}{2}} - 2}{5} \simeq 0.52736,\ u_3 = u_4 u_5 \simeq 0.1726. \tag{1.31}$$

Consequently, solutions are different since their dynamics are governed by different limiters. For example, $\dfrac{\mathrm{d}c_6}{\mathrm{d}t} = u_4$, $\dfrac{\mathrm{d}c_7}{\mathrm{d}t} = u_5$.

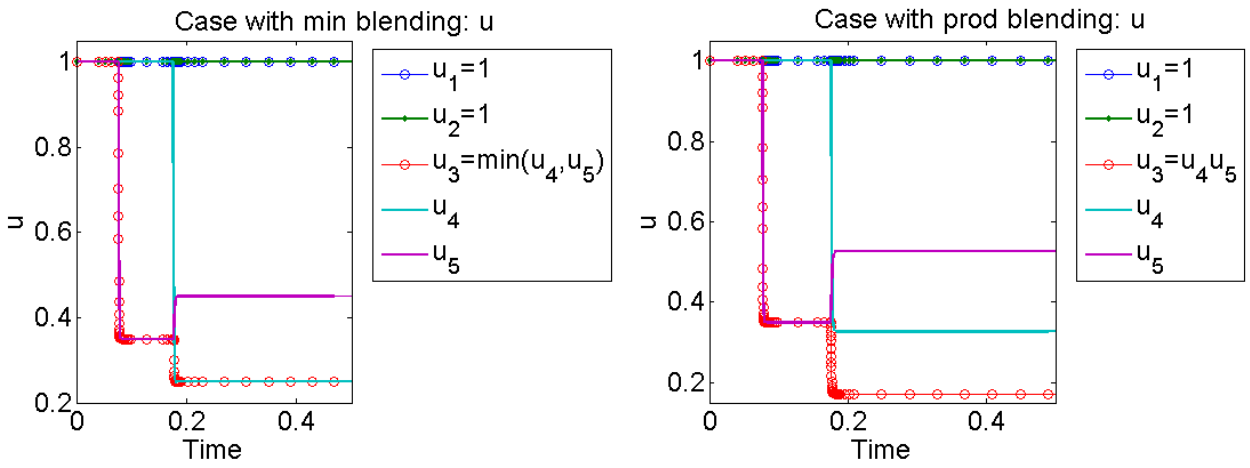


Figure 1.15 – Case with two different blendings. Results with $\epsilon = 0.001$.

Using the limit values of the limiters obtained, system (1.15) has not a unique solution in this case. Indeed for $t \geqslant \max(t_1^*, t_2^*)$, both choices, (1.30) and (1.31), give different solutions of (1.15) satisfying $\frac{\mathrm{d}c_1(t)}{\mathrm{d}t} = 0$ and $\frac{\mathrm{d}c_2(t)}{\mathrm{d}t} = 0$.

Chapter 2

Analysis of a discontinuous ODEs model for multiphase chemical kinetics

2.1 Introduction

In this chapter we present a differential inclusion model describing kinetic reactions. The idea is to consider the right-hand side as a multivalued map, see [38], [39] and [69] for an introduction to this theory. In [2], [12] and [46], authors consider discontinuous right-hand side for kinetic reactions: a discontinuous ODEs formulation in [12] and [46], a complementarity formulation in [46] and a differential inclusion formulation in [2] and [46]. In all of these papers, authors prove existence, positivity and uniqueness of the solution. All of these models assume that reactions cannot contain more than one mineral and every mineral can be present in only one chemical reaction.

In differential inclusion theory, the difficulty is that the right-hand side of the ODE system is no more uniquely defined on the discontinuity surfaces: it is a set-valued function. Hence, an important step of the analysis is to describe the trajectory of a solution. This is done by investigating the properties of the right hand side of the ODE system, which can be considered as a vector field. In [4], [29], [30], [31], [32] and [41], authors characterize different vector field configurations on discontinuity surfaces. In this chapter, we use the result presented in these papers to define the right-hand side of the differential inclusion model for kinetic systems for which only one mineral can disappear.

In chapter 1 (Section 1.4), we defined a differential inclusion system for $\boldsymbol{c}(t) \geqslant 0$:

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} = S\boldsymbol{r}(t,\boldsymbol{c}), \\ \boldsymbol{c}(t) \geqslant 0, \\ \boldsymbol{c}(0) = \boldsymbol{c_0}. \end{cases} \tag{2.1}$$

The reaction rate $\boldsymbol{r}(t,\boldsymbol{c})$ is a continuous function on the domain $[0,T] \times G_+$ where $G_+ = \{\boldsymbol{c} \in \mathbb{R}^{N_a+N_p} \mid \forall i \in \{1,\cdots,N_a+N_p\},\ c_i > 0\}$ is the positive domain. However, $\boldsymbol{r}$ is only defined as a set-valued function on the domain $[0,T] \times \partial G_+$, which is the

domain where at least one mineral is completely dissolved. The objective of this chapter is to better characterize our problem using Filippov's theory [39], in order to use the relevant results and theorems to characterize the solution.
First, we extend the model by defining the reaction rate for strict negative mineral concentrations. This allows us to use Filippov's theory [39] to more precisely define the possible values of $\boldsymbol{r}$ on $[0,T]\times\partial G_+$. With this definition, we can then prove the existence of non-negative solutions to the differential inclusion system.
Second, we will study the different cases resulting from various values of the neighbouring reaction rates around a discontinuity. We precise the cases for which Filippov's original definition leads to uniquely determined reaction rates on $[0,T]\times\partial G_+$ and show that this leads to the existence of a unique non-negative solution to the resulting discontinuous ODEs system that we characterize.

2.2 Differential inclusion model for chemical kinetics

2.2.1 Domains and surfaces of discontinuity

Let $\boldsymbol{c}=(\boldsymbol{c_a},\boldsymbol{c_p})$, where $\boldsymbol{c_a}\in\mathbb{R}^{N_a}$ is the vector of aqueous species concentrations and $\boldsymbol{c_p}\in\mathbb{R}^{N_p}$ is the vector of mineral species concentrations.

We define $\Sigma=\bigcup_{i\in\{1,\cdots,N_p\}}\Sigma_i$, with $\Sigma_i=\{\boldsymbol{c}=(\boldsymbol{c_a},\boldsymbol{c_p})\in\mathbb{R}^{N_a+N_p}\mid c_{p,i}=0\}$, the surface where at least one mineral has zero concentration.

The set Σ divides the space $\mathbb{R}^{N_a+N_p}$ into 2^{N_p} open subspaces G_k, $k\in\{1,\cdots,2^{N_p}\}$, uniquely characterized by the sign of the concentration of the mineral species. It is convenient to use a vector representation $\boldsymbol{e^k}\in\mathbb{R}^{N_p}$ of the domains G_k:

$$e_i^k=\begin{cases}1, & \text{if } \forall \boldsymbol{c}\in G_k,\ c_{p,i}>0,\\ 0, & \text{if } \forall \boldsymbol{c}\in G_k,\ c_{p,i}<0.\end{cases}$$

Let $G:=\bigcup_{k\in\{1,\cdots,2^{N_p}\}}G_k$.

We have $G\cup\Sigma=\mathbb{R}^{N_a+N_p}$, although G and Σ do not form a partition of $\mathbb{R}^{N_a+N_p}$ since the discontinuity surfaces Σ_i intersect. For $\boldsymbol{c}\in\mathbb{R}^{N_a+N_p}$, let $d\geqslant 0$ be the codimension of $\boldsymbol{c}$, that is the number of mineral components equal to 0. The intersections of several discontinuity surfaces Σ_i are characterized by the codimension of each point at the intersection.

A point $\boldsymbol{c}\in\Sigma$ of codimension $d>0$ has 2^d neighbouring domains G_k. We denote by $\mathcal{G}(\boldsymbol{c})$ the set of indices of these domains. A point $\boldsymbol{c}$ in a domain G_k is of codimension 0 and $\mathcal{G}(\boldsymbol{c})=\{k\}$.

Example 2.2.1 (Space decomposition for two minerals)**.** *For a system containing only two minerals, M_1 and M_2, the space is partitioned in four domains G_k, $k\in\{1,2,3,4\}$ and an interface Σ, as shown in Figure 2.1.*

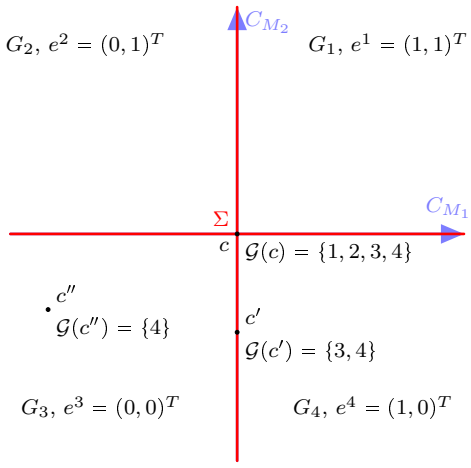


Figure 2.1 – Space decomposition for two minerals

The point $c = (0,0)$ *lies at the intersection of* Σ_1 *and* Σ_2*. It is of codimension* 2 *and has* 4 *neighbouring spaces. The point* $c' = (0,-1.3)$ *is in* Σ_1*, it is of codimension* 1 *and has* 2 *neighbouring spaces.* $c'' = (-2.5,-1)$ *is in* G_3 *and it is of codimension* 0*.*

2.2.2 Extension of reaction rates

The law of mass action in Chapter 1 (Section 1.2.4) is only defined for non-negative aqueous species concentrations. The activity $\boldsymbol{a}$ of an aqueous species i satisfies the following property:

$$\begin{cases} c_{a,i} \geqslant 0 \Rightarrow a_i(\boldsymbol{c_a}) \geqslant 0. \\ c_{a,i} \leqslant 0 \Rightarrow a_i(\boldsymbol{c_a}) = 0. \end{cases} \tag{2.2}$$

We extend this law to negative concentrations by setting their activity to zero in that case:

Definition 2.2.1 (LMA for negative concentrations of aqueous species)**.** $\forall j \in \{1, \cdots, N_r\}$, $\forall \boldsymbol{c} \in \mathbb{R}^{N_a} \times (\mathbb{R}_+^*)^{N_p}$,

$$\tau_j(\boldsymbol{c}) := \alpha_j \prod_{i \in \mathcal{A}_j^-} \big(a_i(\boldsymbol{c_a})\big)^{|S_{i,j}|} - \beta_j \prod_{i \in \mathcal{A}_j^+} \big(a_i(\boldsymbol{c_a})\big)^{|S_{i,j}|}, \tag{2.3}$$

with α_j *and* β_j *are positive constants of reactions.*

We now extend the definition of the reaction rate (1.14) introduced in Chapter 1 to negative concentrations of mineral species.

Definition 2.2.2 (Extended reaction rates in G)**.** *The reaction rates* $\boldsymbol{r} : G \mapsto \mathbb{R}^{N_r}$ *are defined by,* $\forall \boldsymbol{c} \in G$*,*

$$r_j(\boldsymbol{c}) = \min_{i \in \mathcal{M}_j^-} \big(H(c_i)\big)\tau_j^+(\boldsymbol{c}) - \min_{i \in \mathcal{M}_j^+} \big(H(c_i)\big)\tau_j^-(\boldsymbol{c}),\ \forall j \in \{1, \cdots, N_r\}, \tag{2.4}$$

where

$$H(x) = \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{if } x < 0. \end{cases}$$

This definition means that we cancel the reaction rate of any reaction that consumes a mineral whose concentration is negative. In the positive domain $\boldsymbol{c_p} > 0$, all the kinetic reactions proceed as usual : $\boldsymbol{r}(\boldsymbol{c}) = \boldsymbol{\tau}(\boldsymbol{c})$.

The extended reaction rates $\boldsymbol{r}$, defined in (2.4), are uniquely defined in G and are undefined in Σ.

2.2.3 Differential inclusion kinetic model

The idea behind the discontinuous ODE theory developed in [39] and [69] originates from physical considerations: the motion of an ODE solution on a discontinuous surface Σ where the vector field is not provided must depend on the well-defined vector fields in the neighbouring spaces G_k. A natural condition is to define the vector field on Σ as a convex combination of the vector fields in the neighbouring spaces G_k. Filippov's theory then states that the set of admissible values of the vector field in Σ is the convex hull of the limit values of neighbouring vector fields.

The limit values of neighbouring vector fields in Σ are obtained by extending by continuity each function $\boldsymbol{r}_{|\boldsymbol{G_k}} : G_k \mapsto \mathbb{R}^{N_r}$ to $\overline{G_k}$. Let $\boldsymbol{r^k} : \overline{G_k} \mapsto \mathbb{R}^{N_r}$ be this extension such that $\forall k \in \{1, \cdots, 2^{N_p}\}$, $\forall j \in \{1, \cdots, N_r\}$ and $\forall \boldsymbol{c} \in \overline{G_k}$,

$$r_j^k(\boldsymbol{c}) = \big(\min_{i \in \mathcal{M}_j^-} e_i^k\big)\tau_j^+(\boldsymbol{c}) - \big(\min_{i \in \mathcal{M}_j^+} e_i^k\big)\tau_j^-(\boldsymbol{c}). \tag{2.5}$$

We can now define the reaction rates on Σ as the convex hull of the neighbouring $\boldsymbol{r^k}$:

Definition 2.2.3 (Extended reaction rates in $\mathbb{R}^{N_a+N_p}$)**.** *The reaction rates* $\boldsymbol{r} : [0,1]^{2^{N_p}} \times \mathbb{R}^{N_a+N_p} \mapsto \mathbb{R}^{N_r}$ *are defined by,* $\forall \boldsymbol{c} \in \mathbb{R}^{N_a+N_p}$*,*

$$\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}) = \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k \boldsymbol{r^k}(\boldsymbol{c}), \ \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c}), \tag{2.6}$$

with

$$\Lambda(\boldsymbol{c}) = \{\boldsymbol{\lambda} \in [0,1]^{2^{N_p}}, \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k = 1, \ \lambda_i = 0 \ \textit{if} \ i \notin \mathcal{G}(\boldsymbol{c})\}. \tag{2.7}$$

The Filippov set-valued vector field is then $\boldsymbol{F}(\boldsymbol{c}) = \big\{S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}),\ \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c})\big\}$.

Definition 2.2.4 (Kinetic Differential Inclusion system)**.** *We define a kinetic model as a Differential Inclusion system:*

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} \in \big\{S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}),\ \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c})\big\}, \\ \boldsymbol{c}(0) = \boldsymbol{c}_0 \geqslant 0, \end{cases} \tag{2.8}$$

with $\boldsymbol{r}$ *defined by* (2.6)*.*

Remark 2.2.1. *If* $\boldsymbol{c} \in G_k$ *then* $\Lambda(\boldsymbol{c})$ *is reduced to a singleton,* $\Lambda(\boldsymbol{c}) = \{1\}$*, and* $\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}) = \boldsymbol{r^k}(\boldsymbol{c})$ *so that* $\boldsymbol{F}(\boldsymbol{c}) = \{S\boldsymbol{r^k}(\boldsymbol{c})\}$*. In particular,* $\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}) = \boldsymbol{\tau}(\boldsymbol{c})$ *if* $\boldsymbol{c_p} > 0$*.*

We introduce below the definition of a solution to the problem (2.8).

Definition 2.2.5 (Kinetic Differential Inclusion solution)**.** *For a given initial condition* $\boldsymbol{c_0} \geqslant 0$ *and a final time* $T > 0$, $\boldsymbol{c} : [0,T] \mapsto \mathbb{R}^{N_a+N_p}$ *is a solution of* (2.8) *if and only if* $\boldsymbol{c}$ *is an absolutely continuous function and there exists a measurable function* $\boldsymbol{\lambda} : [0,T] \mapsto [0,1]^{2^{N_p}}$ *such that for all* $t \in [0,T]$, $\boldsymbol{\lambda}(t) \in \boldsymbol{\Lambda}(\boldsymbol{c}(t))$ *and*

$$\boldsymbol{c}(t) = \boldsymbol{c}_0 + S\int_0^t \boldsymbol{r}(\boldsymbol{\lambda}(s), \boldsymbol{c}(s))\,\mathrm{d}s. \tag{2.9}$$

We note that Filippov's definition satisfies the conditions of the limited model (1.15) introduced in Chapter 1.

Proposition 2.2.1 (Link between (1.15) and (2.8))**.** *Let* $\boldsymbol{r}$ *be defined by* (2.6). $\forall \boldsymbol{c} \geqslant 0$ *and* $\forall \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c})$,

$$r_j(\boldsymbol{\lambda}, \boldsymbol{c}) = u_j\tau_j^+(\boldsymbol{c}) - v_j\tau_j^-(\boldsymbol{c}),\ \forall j \in \{1, \cdots, N_r\}$$

with $u_j = \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k\big(\min_{i\in\mathcal{M}_j^-} e_i^k\big) \in U_j(\boldsymbol{c})$ *and* $v_j = \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k\big(\min_{i\in\mathcal{M}_j^+} e_i^k\big) \in V_j(\boldsymbol{c})$, *where* U_j *and* V_j *are defined in Chapter 1 (Definition 1.4.1).*

Proof. Let $\boldsymbol{c} \geqslant 0$, let $\boldsymbol{\lambda} \in \Lambda(\boldsymbol{c})$ and let $j \in \{1, \cdots, N_r\}$,

$$\begin{aligned} r_j(\boldsymbol{\lambda}, \boldsymbol{c}) &= \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k r_j^k(\boldsymbol{c}) = \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k\big(\min_{i\in\mathcal{M}_j^-} e_i^k\big)\tau_j^+(\boldsymbol{c}) - \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k\big(\min_{i\in\mathcal{M}_j^+} e_i^k\big)\tau_j^-(\boldsymbol{c}) \\ &= u_j\tau_j^+(\boldsymbol{c}) - v_j\tau_j^-(\boldsymbol{c}), \end{aligned}$$

with $u_j = \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k\big(\min_{i\in\mathcal{M}_j^-} e_i^k\big)$ and $v_j = \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k\big(\min_{i\in\mathcal{M}_j^+} e_i^k\big)$.

It remains to prove that $u_j \in U_j(\boldsymbol{c})$ and $v_j \in V_j(\boldsymbol{c})$. Let us prove it for u_j (a similar proof can be done for v_j):
by construction $u_j \in [0,1]$ because $\sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k = 1$ and $\min_{i\in\mathcal{M}_j^-} e_i^k \in [0,1]$.
If $\boldsymbol{c} \in \mathbb{R}^{N_a+N_p}$ such that $\forall i \in \mathcal{M}_j^-$, $c_i > 0$ then $\forall k \in \mathcal{G}(\boldsymbol{c})$ and $\forall i \in \mathcal{M}_j^-$, $e_i^k = 1$. So, $u_j = \sum_{k\in\mathcal{G}(\boldsymbol{c})} \lambda_k = 1$. □

Remark 2.2.2. *The Differential Inclusion model* (2.8) *is more restrictive than model* (1.15)*. Indeed, the limiters* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *are now linearly dependent of the parameter* $\boldsymbol{\lambda}$.

Example 2.2.2 (Differential Inclusion formulation for the reference example $A \rightleftharpoons M$)**.** *This example is described in Section 1.3.4. Here,* $G_1 := \{\boldsymbol{c} \in \mathbb{R}^2,\ c_M > 0\}$, $G_2 := \{\boldsymbol{c} \in \mathbb{R}^2,\ c_M < 0\}$ *and* $\Sigma := \{\boldsymbol{c} \in \mathbb{R}^2,\ c_M = 0\}$.

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} \in \begin{cases} \{S\tau(\boldsymbol{c})\}, & \text{if } \boldsymbol{c} \in G_1, \\ \{S\tau^+(\boldsymbol{c})\}, & \text{if } \boldsymbol{c} \in G_2, \\ \{S\big(\lambda_1\tau(\boldsymbol{c}) + \lambda_2\tau^+(\boldsymbol{c})\big),\ \lambda_1+\lambda_2 = 1,\ \boldsymbol{\lambda} \in [0,1]^2\}, & \text{if } \boldsymbol{c} \in \Sigma, \end{cases} \\ \boldsymbol{c}(0) = \boldsymbol{c}_0 \geqslant 0. \end{cases} \tag{2.10}$$

We show that (2.8) leads to desirable properties of physical solutions such as existence and non-negativity.

2.2.4 Essential positivity property

Let us show that the extended reaction rates in the negative domain maintain the solution in the positive domain.

Proposition 2.2.2 (Essential Positivity)**.** $\forall \boldsymbol{c} \in \mathbb{R}^{N_p+N_a}$, $\forall i \in \{1, \cdots, N_p + N_a\}$ *such that* $c_i < 0$,

$$(S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}))_i \geqslant 0, \ \forall \boldsymbol{\lambda} \in \Lambda(\boldsymbol{c}).$$

Proof. Let $i \in \{1, \cdots, N_p + N_a\}$, let $\boldsymbol{c} \in \mathbb{R}^{N_a+N_p}$ such that $c_i < 0$.
We recall that for every $\boldsymbol{\lambda} \in \Lambda(\boldsymbol{c})$,

$$\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}) = \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k \boldsymbol{r^k}(\boldsymbol{c}) \text{ with } \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k = 1, \ \lambda_k \in [0, 1], \ \forall k \in \mathcal{G}(\boldsymbol{c}).$$

Hence,

$$(S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}))_i = \sum_{j \in \{1, \cdots, N_r\}} \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k S_{i,j} \Big(\prod_{l \in \mathcal{M}_j^-} e_l^k \Big) \tau_j^+(\boldsymbol{c}) - \sum_{j \in \{1, \cdots, N_r\}} \sum_{k \in \mathcal{G}(\boldsymbol{c})} \lambda_k S_{i,j} \Big(\prod_{l \in \mathcal{M}_j^+} e_l^k \Big) \tau_j^-(\boldsymbol{c}),$$

with $\tau_j(\boldsymbol{c})$ defined in (2.3).
Let $k \in \mathcal{G}(\boldsymbol{c})$ and $j \in \{1, \cdots, N_r\}$. There are two possibilities for the indices i, the species is a mineral or is an aqueous species.

1. Let's consider the aqueous case $i \in \{1, \cdots, N_a\}$,
   1.1. If $i \in \mathcal{A}_j^+$, we have $S_{i,j} > 0$, using the assumption (2.2),

$$c_i < 0 \Rightarrow a_i(\boldsymbol{c_a^+}) = 0 \Rightarrow \tau_j(\boldsymbol{c}) = \alpha_j \prod_{i' \in \mathcal{A}_j^-} \big(a_{i'}(\boldsymbol{c_a^+})\big)^{|S_{i',j}|} \geqslant 0.$$

   Then $\tau_j^+(\boldsymbol{c}) = \tau_j(\boldsymbol{c})$ and $\tau_j^-(\boldsymbol{c}) = 0$.
   So,

$$\begin{cases} \lambda_k S_{i,j} \Big(\prod_{l \in \mathcal{M}_j^-} e_l^k \Big) \tau_j^+(\boldsymbol{c}) \geqslant 0, \\ \lambda_k S_{i,j} \Big(\prod_{l \in \mathcal{M}_j^+} e_l^k \Big) \tau_j^-(\boldsymbol{c}) = 0. \end{cases}$$

   1.2. If $i \in \mathcal{A}_j^-$ then $S_{i,j} < 0$, using the assumption (2.2),

$$c_i < 0 \Rightarrow a_i(\boldsymbol{c_a^+}) = 0 \Rightarrow \tau_j(\boldsymbol{c}) = -\beta_j \prod_{i' \in \mathcal{A}_j^+} \big(a_{i'}(\boldsymbol{c_a^+})\big)^{|S_{i',j}|} \leqslant 0.$$

   Then $-\tau_j^-(\boldsymbol{c}) = \tau_j(\boldsymbol{c})$ and $\tau_j^+(\boldsymbol{c}) = 0$.
   So,

$$\begin{cases} \lambda_k S_{i,j} \Big(\prod_{l \in \mathcal{M}_j^-} e_l^k \Big) \tau_j^+(\boldsymbol{c}) = 0, \\ -\lambda_k S_{i,j} \Big(\prod_{l \in \mathcal{M}_j^+} e_l^k \Big) \tau_j^-(\boldsymbol{c}) \geqslant 0. \end{cases}$$

1.3. Else, $S_{i,j} = 0$ obviously $\lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^-} e_l^k\Big)\tau_j^+(\boldsymbol{c}) = \lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^+} e_l^k\Big)\tau_j^-(\boldsymbol{c}) = 0$.

2. Let's consider the mineral case $i \in \{N_a+1, \cdots, N_p+N_a\}$,
2.1. If $i \in \mathcal{M}_j^+$ then $S_{i,j} > 0$.

$$c_i < 0 \Rightarrow \prod_{l\in\mathcal{M}_j^+} e_l^k = 0.$$

So,

$$\begin{cases} \lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^-} e_l^k\Big)\tau_j^+(\boldsymbol{c}) \geqslant 0, \\ \lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^+} e_l^k\Big)\tau_j^-(\boldsymbol{c}) = 0. \end{cases}$$

2.2. If $i \in \mathcal{M}_j^-$ then $S_{i,j} < 0$.

$$c_i < 0 \Rightarrow \prod_{l\in\mathcal{M}_j^-} e_l^k = 0.$$

So,

$$\begin{cases} \lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^-} e_l^k\Big)\tau_j^+(\boldsymbol{c}) = 0, \\ -\lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^+} e_l^k\Big)\tau_j^-(\boldsymbol{c}) \geqslant 0. \end{cases}$$

2.3. Else, $S_{i,j} = 0$ obviously $\lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^-} e_l^k\Big)\tau_j^+(\boldsymbol{c}) = \lambda_k S_{i,j}\Big(\prod_{l\in\mathcal{M}_j^+} e_l^k\Big)\tau_j^-(\boldsymbol{c}) = 0$.

Summing over all terms, we have $(S\boldsymbol{r}(\boldsymbol{\lambda}, \boldsymbol{c}))_i \geqslant 0$, for every $\boldsymbol{\lambda} \in \Lambda(\boldsymbol{c})$. □

2.2.5 Existence of solutions

As in Chapter 1, let Q be a basis of the kernel of S^T such that $Q^T S = 0$ and let $C_0 = \{\boldsymbol{c} \in (\mathbb{R}_+)^{N_a+N_p},\ Q^T\boldsymbol{c} = Q^T\boldsymbol{c}_0\}$. We have the following result:

Theorem 2.2.1 (Mass conservation and positivity)**.** *Let $\boldsymbol{c} : [0,T] \mapsto \mathbb{R}^{N_a+N_p}$ be a solution of (2.8). Then $\forall t \in [0,T],\ \boldsymbol{c}(t) \in C_0$.*

Proof. Let $\boldsymbol{c} : [0,T] \mapsto \mathbb{R}^{N_a+N_p}$ be a solution of (2.8).

By multiplying (2.9) by Q^T, we get:

$$Q^T\boldsymbol{c}(t) = Q^T\boldsymbol{c}_0 + \int_0^t Q^T S\boldsymbol{r}(\boldsymbol{\lambda}(s), \boldsymbol{c}(s))\,\mathrm{d}s = Q^T\boldsymbol{c}_0.$$

There remains to prove that $\boldsymbol{c}$ is positive. By contradiction, let us assume that there exists $t' \in]0,T]$ and $i \in \{1, \cdots, N_a+N_p\}$ such that $c_i(t') < 0$.
Let $\mathcal{T}_i = \{t \in [0,t'],\ c_i(t) = 0\}$. This set is non-empty because c_i is continuous and $c_i(0) \geqslant 0$. It is closed and bounded because c_i is continuous and $\mathcal{T}_i = [0,t'] \cap c_i^{-1}(\{0\})$.

Therefore, $\mathcal{T}_i$ is an non-empty compact set and admits a maximum element t^*. Using the continuity of c_i, we deduce that $\forall t \in]t^*, t'], \, c_i(t) < 0$.
On the other hand, using (2.9), we have:

$$c_i(t') = c_i(t^*) + \int_{t^*}^{t'} \frac{\mathrm{d}c_i(s)}{\mathrm{d}t}\,\mathrm{d}s = \int_{t^*}^{t'} f_{i_{|\{c_i<0\}}}(\boldsymbol{\lambda}(s), \boldsymbol{c}(s))\,\mathrm{d}s, \text{ where } f_i = (S\boldsymbol{r})_i.$$

Hence, by using the essential positivity Proposition 2.2.2, we get $c_i(t') \geqslant 0$ which contradicts our initial assumption. □

As in Chapter 1, we assume that there exists a conservative variable $\boldsymbol{q}^T\boldsymbol{c}$ defined by a positive vector $\boldsymbol{q}$.

Hypothesis 2.2.1 (Positive conservative variable)**.**

$$\exists \boldsymbol{q} \in Ker(S^T),\ \boldsymbol{q} > 0. \tag{2.11}$$

We enunciate below an existence and boundedness result for the solution of (2.8)

Lemma 2.2.1 (Bounded constraints)**.** *Under Hypothesis 2.2.1, C_0 is bounded.*

Proof. See Lemma 1.2.1. □

Theorem 2.2.2 (Existence)**.** *Under Hypothesis 2.2.1, for all $\boldsymbol{c_0} \in (\mathbb{R}_+)^{N_a+N_p}$, there exists at least one solution $\boldsymbol{c}$ of (2.8). Moreover, $\boldsymbol{c} : [0, +\infty[\to C_0$.*

Proof. Let $\boldsymbol{c_0} \in (\mathbb{R}_+)^{N_a+N_p}$ and let $\boldsymbol{F}$ the right-hand side of (2.8).
Let $\mathcal{B}_0$ be the unit ball of $\mathbb{R}^{N_a+N_p}$. By construction, for all $\boldsymbol{c} \in C_0 + \mathcal{B}_0$, $\boldsymbol{F}(\boldsymbol{c})$ is closed, convex, non-empty and bounded. Moreover, using Lemma 3 in [39] page 67, $\boldsymbol{F}$ is upper semicontinuous. The assumptions of Theorem 1 in [39] page 77 are then satisfied.
Therefore, there exists $T \in [0, +\infty[$ and $\boldsymbol{c} : [0, T] \to C_0 + \mathcal{B}_0$ an absolutely continuous function solution of

$$\frac{\mathrm{d}\boldsymbol{c}(t)}{\mathrm{d}t} \in \boldsymbol{F}_{|\boldsymbol{C_0}+\boldsymbol{\mathcal{B}_0}}(\boldsymbol{c}(t)), \quad \boldsymbol{c}(0) = \boldsymbol{c_0},$$

with $T \geqslant \dfrac{0.5}{\sup\limits_{\mathcal{B}(\boldsymbol{c_0},0.5)} |\boldsymbol{F}(\boldsymbol{c})|}$.
Combining the mass conservation property and the positivity proved in Theorem 2.2.1, every solution of (2.8) starting in C_0 remains in C_0. Hence, $\boldsymbol{c} : [0, T] \to C_0$.
Moreover, we have $T \geqslant \dfrac{0.5}{\sup\limits_{C_0+\mathcal{B}_0} |\boldsymbol{F}(\boldsymbol{c})|} > 0$. Using that the existence time T is bigger than a strict positive constant and by repeating the same process considering $\boldsymbol{c}(T)$ as the initial time (by construction the set C_0 is the same), we prove that there exists $\boldsymbol{c} : [0, 2T] \to C_0$ solution of (2.8). This concludes the proof. □

2.3 Uniqueness of the solution when only one mineral disappears

We study the case where only one mineral can be dissolved. Without loss of generality, we denote M_1 this mineral and $c_1 = c_{p,1}$ the concentration of M_1. Let $G_1 = \{(\boldsymbol{c_a}, \boldsymbol{c_p}),\ \boldsymbol{c_a} \geqslant 0,\ \boldsymbol{c_p} > 0\}$, $G_2 = \{(\boldsymbol{c_a}, \boldsymbol{c_p}),\ \boldsymbol{c_a} \geqslant 0,\ c_{p,1} < 0,\ \boldsymbol{c_{p,i}} > 0,\ i \geqslant 2\}$ and $\Sigma_1^+ = \{(\boldsymbol{c_a}, \boldsymbol{c_p}),\ \boldsymbol{c_a} \geqslant 0,\ c_{p,1} = 0,\ c_{p,i} > 0,\ i \geqslant 2\}$. Thus any $\boldsymbol{c} \in \Sigma_1^+$ is of codimension 1.

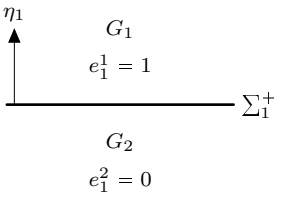


Figure 2.2 – Space decomposition in codimension 1

Let $\boldsymbol{f^1}$ (respectively $\boldsymbol{f^2}$) be the extended right-hand side of (2.8) in $\overline{G_1}$ (respectively $\overline{G_2}$).

The set-valued function $\boldsymbol{F}$ is defined as follows

$$\boldsymbol{F}(\boldsymbol{c}) = \begin{cases} \{\boldsymbol{f^1}(\boldsymbol{c})\}, & \boldsymbol{c} \in G_1, \\ \{\boldsymbol{f^2}(\boldsymbol{c})\}, & \boldsymbol{c} \in G_2, \\ \{\lambda \boldsymbol{f^1}(\boldsymbol{c}) + (1-\lambda)\boldsymbol{f^2}(\boldsymbol{c}),\ \lambda \in [0,1]\}, & \boldsymbol{c} \in \Sigma_1^+, \end{cases} \tag{2.12}$$

where $\boldsymbol{f^1}(\boldsymbol{c}) = S\boldsymbol{\tau}(\boldsymbol{c})$ and $\boldsymbol{f^2}(\boldsymbol{c}) = S\boldsymbol{r^2}(\boldsymbol{c})$.

The normal components f_1^1 and f_1^2 are given by:

$$\begin{cases} f_1^1(\boldsymbol{c}) = \displaystyle\sum_{j \in \{1,\cdots,N_r\}} S_{1,j}\tau_j(\boldsymbol{c}), & \boldsymbol{c} \in \overline{G_1} \\ f_1^2(\boldsymbol{c}) = \displaystyle\sum_{j \in \{1,\cdots,N_r\}} \left(S_{1,j}^-\tau_j^-(\boldsymbol{c}) + S_{1,j}^+\tau_j^+(\boldsymbol{c})\right), & \boldsymbol{c} \in \overline{G_2}. \end{cases} \tag{2.13}$$

Before describing the possible configurations of the kinetic fields we present the following result

Proposition 2.3.1 (Partial attractiveness of kinetic fields)**.** *The vector field satisfies the following assertions:*

— $\forall \boldsymbol{c} \in G_2 \cup \Sigma_1^+,\ f_1^2(\boldsymbol{c}) \geqslant 0$,

— $\forall \boldsymbol{c} \in \Sigma_1^+,\ f_1^2(\boldsymbol{c}) = 0 \Rightarrow f_1^1(\boldsymbol{c}) \leqslant 0$.

Proof. By sum of positive terms and using Proposition 2.2.2, the assertion $\forall \boldsymbol{c} \in G_2 \cup \Sigma_1^+,\ f_1^2(\boldsymbol{c}) \geqslant 0$ holds.

Let us prove the second assertion. Let $\boldsymbol{c} \in \Sigma_1^+$ such that $f_1^2(\boldsymbol{c}) = 0$,

$$f_1^1(\boldsymbol{c}) = f_1^1(\boldsymbol{c}) - f_1^2(\boldsymbol{c}) = -\sum_{j \in \{1,\cdots,N_r\}} \left(S_{1,j}^-\tau_j^+(\boldsymbol{c}) + S_{1,j}^+\tau_j^-(\boldsymbol{c})\right) \leqslant 0.$$

That concludes the proof. □

Let $\boldsymbol{c} \in \Sigma_1^+$, there are three possible configurations for the kinetic fields:

— Attractive: $[f_1^1(\boldsymbol{c}) < 0$ (and $f_1^2(\boldsymbol{c}) \geqslant 0)]$ or $[f_1^1(\boldsymbol{c}) = 0$ and $f_1^2(\boldsymbol{c}) > 0]$.
— Transverse: $f_1^1(\boldsymbol{c}) > 0$ (and $f_1^2(\boldsymbol{c}) > 0$, thanks to Proposition 2.3.1).
— Tangent: $f_1^1(\boldsymbol{c}) = 0$ and $f_1^2(\boldsymbol{c}) = 0$.

In order to prove the uniqueness of the solution of (2.8) in the case of pointwise attractive or transverse field in codimension 1, we recall a classical result proved by Filippov in [39].

Theorem 2.3.1 (Theorem 2 page 110 in [39]). *Let $G \subset \mathbb{R}^n$ be separated by a smooth surface $\Sigma \in \mathcal{C}^2$, with η a normal vector, into domains G_1 and G_2. Let f and $\frac{\partial f}{\partial x_i}$, $i = 1, \cdots, n$ be continuous in domains $[0, t_f] \times G_1$ and $[0, t_f] \times G_2$ up to the boundary. Let the function $f^1 - f^2 \in \mathcal{C}^1$ in $[0, t_f] \times \Sigma$.*
If for all $(t, x) \in [0, t_f] \times \Sigma$ at least one of the inequalities $f^1(t, x) \cdot \eta < 0$ or $f^2(t, x) \cdot \eta > 0$ is fulfilled, the right uniqueness of the solution in G of the following problem occurs,

$$\begin{cases} \frac{\mathrm{d}x(t)}{\mathrm{d}t} \in F(t, x(t)), \\ x(0) = x_0 \in G, \end{cases} \tag{2.14}$$

$$\text{with } F(t, x(t)) = \begin{cases} \{f^1(t, x(t))\}, & x(t) \in G_1, \\ \{f^2(t, x(t))\}, & x(t) \in G_2, \\ \{\lambda f^1(t, x(t)) + (1 - \lambda) f^2(t, x(t)),\ \lambda \in [0, 1]\}, & x(t) \in \Sigma. \end{cases}$$

The following uniqueness result is a direct consequence of Theorem 2.3.1.

Corollary 2.3.1 (Uniqueness). *Let $G = \{(\boldsymbol{c_a}, \boldsymbol{c_p}) \geqslant 0,\ c_{p,i} > 0,\ \forall i \geqslant 2\}$ and let $\boldsymbol{c_0} \in G$. If for all $\boldsymbol{c}$ in Σ_1^+, $f_1^1(\boldsymbol{c}) \neq 0$ or $f_1^2(\boldsymbol{c}) \neq 0$ then the system (2.8) has a unique solution in G.*

Proof. Let $\boldsymbol{c} \in \Sigma_1^+$ such that $f_1^1(\boldsymbol{c}) \neq 0$ or $f_1^2(\boldsymbol{c}) \neq 0$ then the field in $\boldsymbol{c}$ is in one of the two configurations:
— $[f_1^1(\boldsymbol{c}) < 0$ and $f_1^2(\boldsymbol{c} \geqslant 0)]$ or $[f_1^1(\boldsymbol{c}) = 0$ and $f_1^2(\boldsymbol{c}) > 0]$,
— $f_1^1(\boldsymbol{c}) > 0$ and $f_1^2(\boldsymbol{c}) > 0$.
Hence $\forall \boldsymbol{c} \in \Sigma_1^+$, $f_1^1(\boldsymbol{c}) < 0$ or $f_1^2(\boldsymbol{c}) > 0$.
The result holds using Theorem 2.3.1 in $\tilde{G} = \{(\boldsymbol{c_a}, \boldsymbol{c_p}) \in \mathbb{R}^{N_a + N_p},\ c_{p,i} > 0,\ \forall i \geqslant 2\}$ combined with Theorem 2.2.1. □

Remark 2.3.1. *In particular, Corollary 2.3.1 shows the uniqueness in the case where the field is attractive (or tangent) for all the points in Σ_1^+.*

The following result shows the positivity of mineral, other that M_1, components.

Proposition 2.3.2 (Positive mineral concentrations). *Let $t^* \in [0, T[$ and $\boldsymbol{c}$ be a solution of (2.8) such that $\boldsymbol{c}(t^*) \in \Sigma_1^+$. There exists $t' \in]t^*, T]$ such that for all $t \in [t^*, t']$, $c_{p,i}(t) > 0$, $\forall i \geqslant 2$.*

Proof. The proof is done using the continuity of $\boldsymbol{c}$. □

In the next sections we describe in more details the three possible configurations for the kinetic fields. In particular, we describe the trajectories of the solutions.

2.3.1 Attractive field

The attractive field configuration occurs if the field in G_1 points towards the discontinuity surface Σ_1^+ with at most one of both vector fields equal to 0. See Figure 2.3 for an illustration.

Definition 2.3.1 (Attractive field)**.** *A kinetic discontinuous field $\boldsymbol{f}$ is attractive at $\boldsymbol{c} \in \Sigma_1^+$ if and only if*

$$[f_1^1(\boldsymbol{c}) < 0 \text{ (and } f_1^2(\boldsymbol{c}) \geqslant 0)] \text{ or } [f_1^1(\boldsymbol{c}) = 0 \text{ and } f_1^2(\boldsymbol{c}) > 0]. \tag{2.15}$$

Remark 2.3.2. *For kinetics right-hand side, (2.15) is equivalent to $f^1(\boldsymbol{c})f^2(\boldsymbol{c}) \leqslant 0$ and $[f^1(\boldsymbol{c}) \neq 0$ or $f^2(\boldsymbol{c}) \neq 0]$*

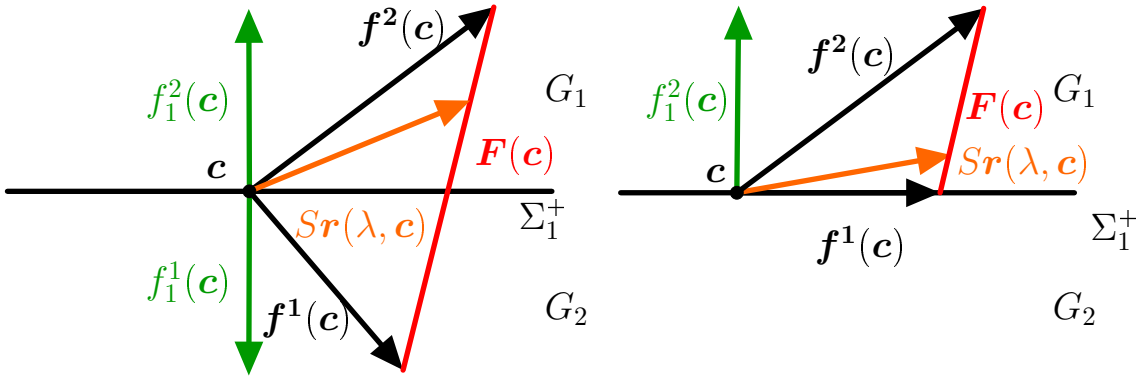


Figure 2.3 – Field on a discontinuity surface in the attractive case. Left: strict attractive case. Right: attractive case with a tangent vector field $\boldsymbol{f^1}$.

We study the trajectory of the unique solution in the attractive case. Let introduce the following assumption.

Hypothesis 2.3.1 (Attractive neighbourhood)**.** *Let $\boldsymbol{c} \in \Sigma_1^+$ such that the field is attractive. There exists $\mathcal{V}$ an open neighbourhood of $\boldsymbol{c}$ such that $\forall \boldsymbol{c}^* \in \overline{G_1} \cap \mathcal{V}$, $f_1^1(\boldsymbol{c}^*) \leqslant 0$.*

The following result ensures that in the attractive case, the solution remains on Σ_1^+.

Theorem 2.3.2 (Attractive sliding motion)**.** *Let $\boldsymbol{c} : [0, T] \mapsto C_0$ be a solution of (2.8). Suppose that $\exists t^* \in [0, T[$ such that $\boldsymbol{c}(t^*) \in \Sigma_1^+$, the attractive conditions are satisfied at $\boldsymbol{c}(t^*)$ and we assume that Hypothesis 2.3.1 holds true. Then there exists $t' > t^*$ such that for all $t \in [t^*, t']$, the trajectory remains on Σ_1^+.*

Proof. Let $\boldsymbol{c} : [0, T] \mapsto C_0$ be a solution of (2.8). Let $t^* \in [0, T[$ such that $\boldsymbol{c}(t^*) \in \Sigma_1^+$ and the attractive conditions are satisfied at $\boldsymbol{c}(t^*)$.
Using Proposition 2.3.2, we know that all minerals components $c_{p,i}$, $\forall i \geqslant 2$, are strictly positive. It remains to prove that there exists $t' > t^*$ such that for all $t \in [t^*, t']$, $c_1(t) = 0$. Using Hypothesis 2.3.1, there exists $\mathcal{V}$ an open neighbourhood of $\boldsymbol{c}(t^*)$ such that $\forall \boldsymbol{c}^* \in \overline{G_1} \cap \mathcal{V}$, $f_1^1(\boldsymbol{c}^*) \leqslant 0$. Using the continuity and the non-negativity of $\boldsymbol{c}$, there exists $t' > t^*$

such that $\forall s \in [t^*, t'],\ \boldsymbol{c}(s) \in \mathcal{V} \cap \overline{G_1}$. Hence, $\forall s \in [t^*, t'],\ f_1^1(\boldsymbol{c}(s)) \leqslant 0$.
Let's prove by contradiction that $\forall s \in [t^*, t'],\ c_1(s) = 0$.
We assume that there exists $z \in [t^*, t']$ such that $c_1(z) > 0$. Let $A = \{t \in [t^*, z],\ c_1(t) = 0\}$ be a non-empty and compact set and let z_1 be its maximal element. By definition $z_1 < z$ and we have:

$$c_1(z) = c_1(z_1) + \int_{z_1}^{z} f_1^1(s)\ \mathrm{d}s \leqslant 0,$$

which is a contradiction and that concludes the proof. □

We then determine a unique sliding field on Σ_1^+ in the attractive case.

Corollary 2.3.2 (Sliding field)**.** *Under the same assumptions of Theorem 2.3.2, there exists $t' > t^*$ such that for almost every $t \in [t^*, t']$, $\boldsymbol{F}(\boldsymbol{c}(t)) = \left\{\boldsymbol{f}(\lambda(t), \boldsymbol{c}(t)),\ \lambda(t) = \dfrac{f_1^2(\boldsymbol{c}(t))}{f_1^2(\boldsymbol{c}(t)) - f_1^1(\boldsymbol{c}(t))}\right\}$ is a singleton.*

Proof. Using Theorem 2.3.2, there exists $t' > t^*$ such that for almost every $t \in [t^*, t']$, $F_1(\boldsymbol{c}(t)) = \{0\}$. So, for almost every $t \in [t^*, t']$, the right-hand side satisfies the following system:

$$\begin{cases} f_1(\lambda(t), \boldsymbol{c}(t)) = 0, \\ \boldsymbol{f}(\lambda(t), \boldsymbol{c}(t)) = \lambda(t)\boldsymbol{f}^1(\boldsymbol{c}(t)) + (1 - \lambda(t))\boldsymbol{f}^2(\boldsymbol{c}(t)), \\ \lambda(t) \in [0, 1], \end{cases} \tag{2.16}$$

which has a unique solution

$$\lambda(t) = \frac{f_1^2(\boldsymbol{c}(t))}{f_1^2(\boldsymbol{c}(t)) - f_1^1(\boldsymbol{c}(t))}.$$

By hypothesis $f_1^1(\boldsymbol{c}(t)) \leqslant 0$, $f_1^2(\boldsymbol{c}(t)) \geqslant 0$ and $f_1^1(\boldsymbol{c}(t)) - f_1^2(\boldsymbol{c}(t)) \neq 0$ so $\lambda(t) \in [0, 1]$. □

Figure 2.4 illustrates the attractive field in codimension 1. The two vector fields $\boldsymbol{f^1}$ and $\boldsymbol{f^2}$ satisfy the attractive condition. We draw in green the normal component of both fields and in red the convex hull $\boldsymbol{F}$ and in orange the sliding value $\boldsymbol{f}^{\Sigma_1^+}$ of the right-hand side of (2.16).

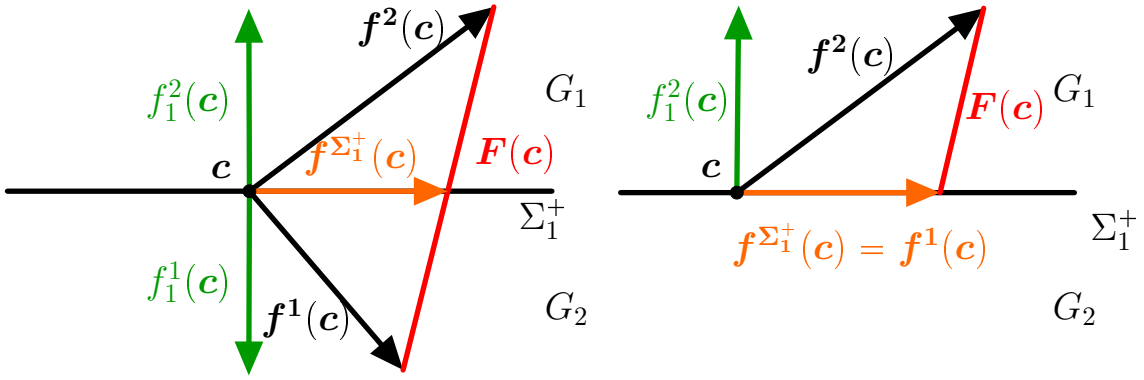


Figure 2.4 – Right-hand side in the attractive case at $\boldsymbol{c}$ of codimension 1

We discuss below the relevance of Hypothesis 2.3.1. Indeed it is important to emphasize that this hypothesis can be proved in the case $\boldsymbol{c} \in \Sigma_1^+$ if $f_1^1(\boldsymbol{c}) < 0$.

Proposition 2.3.3 (Attractive field in a neighbourhood in the case $f_1^1 < 0$)**.** *If $\boldsymbol{f}$ is an attractive field at $\boldsymbol{c} \in \Sigma_1^+$ such that $f_1^1(\boldsymbol{c}) < 0$ then there exists a neighbourhood $\mathcal{V}$ of $\boldsymbol{c}$ such that*

$$\forall \boldsymbol{c}^* \in \overline{G_1} \cap \mathcal{V},\ f_1^1(\boldsymbol{c}^*) < 0.$$

So, Hypothesis 2.3.1 holds true.

Proof. Since f_1^1 is continuous in $\overline{G_1}$ we get the desired result. □

The two examples below (Example 2.3.1 and 2.3.1) correspond to attractive situations of the case $M_1M_2M_3$ described in Chapter 1, when M_2 disappears first. So here the kinetic field is discontinuous on Σ_2^+ defined by

$$\boldsymbol{f^1}(\boldsymbol{c}) = \begin{pmatrix} -\tau_1 \\ \tau_1 - \tau_2 \\ \tau_2 \end{pmatrix},\ \boldsymbol{f^2}(\boldsymbol{c}) = \begin{pmatrix} -\tau_1 \\ \tau_1 \\ 0 \end{pmatrix},$$

with $\tau_1 > 0$ and $\tau_2 > 0$.
Using the analysis developed in this chapter, we get the same value for the right-hand side on the discontinuity surface.

Example 2.3.1 (case $M_1M_2M_3$ when M_2 disappears first in case $f_1^1 < 0$)**.** *Let consider the case $\tau_1 - \tau_2 < 0$. In this case the field is attractive because $f_2^1(\boldsymbol{c}) < 0$. Then Corollary 2.3.2 yields $\lambda = \dfrac{\tau_1}{\tau_2}$ and the unique possible value of the right-hand side $\boldsymbol{F}$ on $\Sigma_2^+ = \{c_2 = 0,\ c_1 > 0\}$ is $\boldsymbol{f^{\Sigma_2^+}}$ defined as follow,*

$$\boldsymbol{f^{\Sigma_2^+}} = \begin{pmatrix} -\tau_1 \\ 0 \\ \tau_1 \end{pmatrix}.$$

We recover the limited right-hand side defined in Chapter 1.

In the case where $f_1^1(\boldsymbol{c}) = 0$ and $f_1^2(\boldsymbol{c}) > 0$ at $\boldsymbol{c} \in \Sigma_1^+$, Hypothesis 2.3.1 is not necessarily satisfied: a difficult case arises when $f_1^1(\boldsymbol{c}) > 0$ for $\boldsymbol{c} \in G_1$ but $f_1^1(\boldsymbol{c}) = 0$ for $\boldsymbol{c} \in \Sigma_1^+$, i.e. only the limit of the vector field is tangent to Σ_1^+. The proof of Theorem 2.3.2 doesn't apply and we can't determine if the solution remains on Σ_1^+ or enters G_1. Nevertheless, there exist chemical systems with tangent vector fields satisfying Hypothesis 2.3.1:

Example 2.3.2 (case $M_1M_2M_3$ when M_2 disappears first in the case $f_1^1 = 0$ and $f_1^2 > 0$)**.** *Let consider the case $\tau_1 - \tau_2 = 0$. In this case the field is attractive because $f_2^1(\boldsymbol{c}) = 0$ and $f_2^2(\boldsymbol{c}) > 0$. Hypothesis 2.3.1 holds true. Then Corollary 2.3.2, in this context, yields $\lambda = 1$ and the unique possible value of the right-hand side $\boldsymbol{F}$ on $\Sigma_2^+ = \{c_2 = 0,\ c_1 > 0\}$ is $\boldsymbol{f^{\Sigma_2^+}}$ defined as follow,*

$$\boldsymbol{f^{\Sigma_2^+}} = \begin{pmatrix} -\tau_1 \\ 0 \\ \tau_2 \end{pmatrix}.$$

2.3.2 Transverse field

The transverse field configuration is defined by the two neighbouring vector fields pointing in the same direction. See Figure 2.5 for an illustration.

Definition 2.3.2 (Transverse field)**.** *A kinetic discontinuous field* $\boldsymbol{f}$ *is transverse at* $\boldsymbol{c} \in \Sigma_1^+$ *if and only if*

$$f_1^1(\boldsymbol{c}) > 0. \tag{2.17}$$

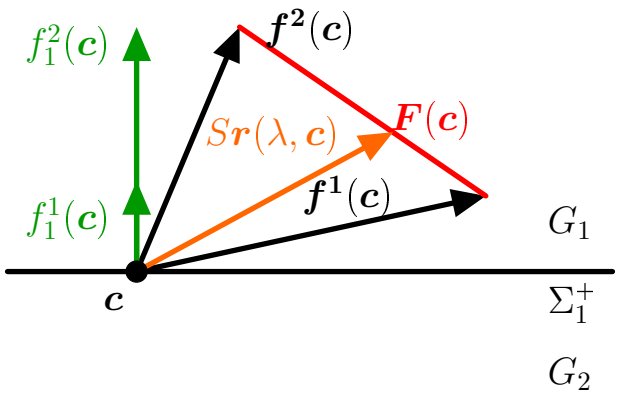


Figure 2.5 – Field on a discontinuity surface in the transverse case

Lemma 2.3.1 (Transverse field in a neighbourhood)**.** *If* $\boldsymbol{f}$ *is a transverse field at* $\boldsymbol{c} \in \Sigma_1^+$ *then there exists a neighbourhood* $\mathcal{V}$ *of* $\boldsymbol{c}$ *such that*

- $\forall \boldsymbol{c}^* \in \mathcal{V}$, $c_{p,i}^* > 0$, $\forall i \geqslant 2$,
- $\forall \boldsymbol{c}^* \in \Sigma_1^+ \cap \mathcal{V}$, $\forall \lambda \in [0,1]$, $\lambda f_1^1(\boldsymbol{c}^*) + (1-\lambda) f_1^2(\boldsymbol{c}^*) > 0$.
- $\forall \boldsymbol{c}^* \in G_1 \cap \mathcal{V}$, $f_1^1(\boldsymbol{c}^*) > 0$.

Proof. Using Proposition 2.3.1, we get $f_1^2(\boldsymbol{c}) > 0$. By continuity, as in the proof of Proposition 2.3.3, we conclude the proof. □

All possible values of the vector field $\boldsymbol{F}$ point toward G_1. Therefore, it seems clear that any solution will leave Σ_1^+ to enter G_1. The result below shows that remarkably, the choice of the vector field on the discontinuity has no impact on the solution.

Theorem 2.3.3 (Solutions leaving the surface)**.** *Let* $\boldsymbol{c}$ *a solution of* (2.8)*. Let* $t^* \in [0, T[$ *such that* $\boldsymbol{c}(t^*) \in \Sigma_1^+$ *and* $f_1^1(\boldsymbol{c}(t^*)) > 0$.
There exists $t' \in]t^*, T[$ *such that* $\forall s \in]t^*, t']$, $c_1(s) > 0$.

Proof. Let $\boldsymbol{c}$ be a solution of (2.8) and let $t^* \in [0, T[$ such that $\boldsymbol{c}(t^*) \in \Sigma_1^+$ and $f_1^1(\boldsymbol{c}(t^*)) > 0$.
There exists $t' \in]t^*, T[$ such that $\forall s \in [t^*, t']$, $\boldsymbol{c}(s) \in \mathcal{V}$, neighbourhood of $\boldsymbol{c}(t^*)$ defined in Lemma 2.3.1. Moreover, Using Theorem 2.2.1, $\forall s \in [t^*, t']$, $\boldsymbol{c}(s) \in (G_1 \cup \Sigma_1^+) \cap \mathcal{V}$.
Let $t \in]t^*, t']$. Using Definition 2.2.5, there exists a function $\lambda : [0, T] \mapsto [0, 1]$, satisfying

- $\forall s \in [t^*, t']$ such that $\boldsymbol{c}(s) \in G_1$, $\lambda(s) = 1$,
- $c_1(t) = c_1(t^*) + \displaystyle\int_{t^*}^{t} \lambda(s) f_1^1(\boldsymbol{c}(s)) + (1-\lambda(s)) f_1^2(\boldsymbol{c}(s))\, \mathrm{d}s$.

So, using Lemma 2.3.1, $\forall t \in]t^*, t'], c_1(t) > 0$. □

The following example shows a chemical system in this situation.

Example 2.3.3 (case $M_1M_2M_3$ when M_2 start at a 0 concentration in the transverse case)**.** *In the case $M_1M_2M_3$ when M_2 is the only one to disappear, both reactions have a positive orientation, $\tau_1 > 0$, $\tau_2 > 0$ and $\tau_1 - \tau_2 > 0$ and the concentration value of M_2 is equal to zero, $c_{0,2} = 0$ (this corresponds chemically to consider M_2 absent at the start of chemical process), we get,*

$$\boldsymbol{f^1}(\boldsymbol{c}) = \begin{pmatrix} -\tau_1 \\ \tau_1 - \tau_2 \\ \tau_2 \end{pmatrix}, \ \boldsymbol{f^2}(\boldsymbol{c}) = \begin{pmatrix} -\tau_1 \\ \tau_1 \\ 0 \end{pmatrix}.$$

In this case the field is transverse because $f_2^1(\boldsymbol{c}) > 0$ and $f_2^2(\boldsymbol{c}) > 0$. Thus, for all possible choices of $\boldsymbol{F}$ at the initial point, the solution are governed by $\boldsymbol{f^1}$ after the initial time (until the disappearance of M_1):

$$\boldsymbol{c}(t) = \boldsymbol{c_0} + \int_0^t \boldsymbol{f^1}(\boldsymbol{c}(s))\,\mathrm{d}s, \ \forall t > 0 \text{ and } c_1(t) > 0.$$

2.3.3 Tangent field

The tangent field configuration corresponds to the case where the first component of both fields are equal to zero on a point of the discontinuity surface Σ_1^+. See Figure 2.6 for an illustration.

Definition 2.3.3 (Tangent field)**.** *A kinetic discontinuous field $\boldsymbol{f}$ is tangent at $\boldsymbol{c} \in \Sigma_1^+$ if and only if*

$$f_1^1(\boldsymbol{c}) = 0 \text{ and } f_1^2(\boldsymbol{c}) = 0. \tag{2.18}$$

A direct consequence of this definition is $F_1(\boldsymbol{c}) = \{0\}$ for all $\boldsymbol{c}$ satisfying the tangent condition.

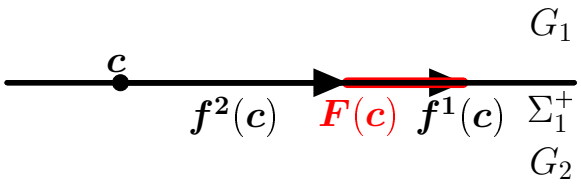


Figure 2.6 – Field on a discontinuity surface in the tangent case

In the tangent case we have the following Proposition.

Proposition 2.3.4 (Tangent conditions)**.** *Let $\boldsymbol{c} \in \Sigma_1^+$ such that $\boldsymbol{f}$ is tangent in $\boldsymbol{c}$. Then the reaction rates of reactions containing M_1 are equal to zero, i.e. $\forall j \in \{1, \cdots, N_r\}$, if $S_{1,j} \neq 0$ then $\tau_j(\boldsymbol{c}) = 0$.*

Proof. Let $\boldsymbol{c} \in \Sigma_1^+$ such that $\boldsymbol{f}$ is tangent in $\boldsymbol{c}$, i.e.

$$\begin{cases} f_1^1(\boldsymbol{c}) = \displaystyle\sum_{j\in\{1,\cdots,N_r\}} S_{1,j}\tau_j(\boldsymbol{c}) = 0, \\ f_1^2(\boldsymbol{c}) = \displaystyle\sum_{j\in\{1,\cdots,N_r\}} \Big(S_{1,j}^-\tau_j^-(\boldsymbol{c}) + S_{1,j}^+\tau_j^+(\boldsymbol{c})\Big) = 0. \end{cases}$$

Hence,

$$f_1^2(\boldsymbol{c}) = 0 \Rightarrow \forall j \in \{1,\cdots,N_r\},\ S_{1,j}^+\tau_j^+(\boldsymbol{c}) = 0 \text{ and } S_{1,j}^-\tau_j^-(\boldsymbol{c}) = 0,$$

using that

$$f_1^1(\boldsymbol{c}) = \sum_{j\in\{1,\cdots,N_r\}} \Big(S_{1,j}^-\tau_j^-(\boldsymbol{c}) + S_{1,j}^+\tau_j^+(\boldsymbol{c})\Big) - \sum_{j\in\{1,\cdots,N_r\}} \Big(S_{1,j}^+\tau_j^-(\boldsymbol{c}) + S_{1,j}^-\tau_j^+(\boldsymbol{c})\Big),$$

$$f_1^1(\boldsymbol{c}) = 0 \text{ and } f_1^2(\boldsymbol{c}) = 0 \Rightarrow \forall j \in \{1,\cdots,N_r\},\ S_{1,j}^+\tau_j^-(\boldsymbol{c}) = 0 \text{ and } S_{1,j}^-\tau_j^+(\boldsymbol{c}) = 0.$$

Gathering previous properties we get $\forall j \in \{1,\cdots,N_r\},\ S_{1,j}\tau_j(\boldsymbol{c}) = 0$. □

Corollary 2.3.3 (Continuity of the field in the tangent case)**.** *Let $\boldsymbol{c} \in \Sigma_1^+$ such that $\boldsymbol{f}$ is tangent in $\boldsymbol{c}$. Then, $\boldsymbol{F}(\boldsymbol{c}) = \{\boldsymbol{f^1}(\boldsymbol{c})\} = \{\boldsymbol{f^2}(\boldsymbol{c})\}$.*

Proof. Let $\boldsymbol{c} \in \Sigma_1^+$ such that $\boldsymbol{f}$ is tangent in $\boldsymbol{c}$. Let $k \in \{2,\cdots,N_a+N_p\}$.

$$\begin{cases} f_k^1(\boldsymbol{c}) = \displaystyle\sum_{\substack{j\in\{1,\cdots,N_r\}\\ S_{1,j}=0}} S_{k,j}\tau_j(\boldsymbol{c}) + \sum_{\substack{j\in\{1,\cdots,N_r\}\\ S_{1,j}\neq 0}} S_{k,j}\tau_j(\boldsymbol{c}), \\ f_k^2(\boldsymbol{c}) = \displaystyle\sum_{\substack{j\in\{1,\cdots,N_r\}\\ S_{1,j}=0}} S_{k,j}\tau_j(\boldsymbol{c}) + \sum_{\substack{j\in\{1,\cdots,N_r\}\\ S_{1,j}>0}} S_{k,j}\tau_j^+(\boldsymbol{c}) - \sum_{\substack{j\in\{1,\cdots,N_r\}\\ S_{1,j}<0}} S_{k,j}\tau_j^-(\boldsymbol{c}). \end{cases}$$

Using Proposition 2.3.4, $\forall j \in \{1,\cdots,N_r\}$ such that $S_{1,j} \neq 0$, $\tau_j(\boldsymbol{c}) = 0$.
So,

$$f_k^1(\boldsymbol{c}) = f_k^2(\boldsymbol{c}) = \sum_{\substack{j\in\{1,\cdots,N_r\}\\ S_{1,j}=0}} S_{k,j}\tau_j(\boldsymbol{c}).$$

This concludes the proof. □

In the tangent case, the assumptions of Theorem 2.3.1 are not satisfied. However, the previous corollary means that the field $\boldsymbol{f}$ is actually continuous in the tangent case. Therefore, if we assume that these tangent conditions hold true in an open neighbourhood of a point $\boldsymbol{c} \in \Sigma_1^+$, then we have a unique solution of (2.8) due to the Lipschitz continuity of $\boldsymbol{f}$ in that neighbourhood. Unfortunately, this assumption does not always hold as shown in the following example:

Example 2.3.4 (Reference example $A \rightleftharpoons M$ in tangent configuration)**.** *Let denote $\tau = \alpha c_a - \beta$ the reaction rate with $\alpha > 0$ and $\beta > 0$. We remind the discontinuous system in this case:*

$$\begin{cases} \dfrac{\mathrm{d}c_a}{\mathrm{d}t} = -\big((\alpha c_a - \beta)^+ - H(c_p)(\alpha c_a - \beta)^-\big), \\ \dfrac{\mathrm{d}c_p}{\mathrm{d}t} = (\alpha c_a - \beta)^+ - H(c_p)(\alpha c_a - \beta)^-. \end{cases}$$

So, the values of the discontinuous right-hand side in all domains are:

— $c_p > 0$ *then* $\boldsymbol{f^1}(\boldsymbol{c}) = \begin{pmatrix} -(\alpha c_a - \beta)^+ + (\alpha c_a - \beta)^- \\ (\alpha c_a - \beta)^+ - (\alpha c_a - \beta)^- \end{pmatrix} = \begin{pmatrix} -(\alpha c_a - \beta) \\ (\alpha c_a - \beta) \end{pmatrix}$,

— $c_p < 0$ *then* $\boldsymbol{f^2}(\boldsymbol{c}) = \begin{pmatrix} -(\alpha c_a - \beta)^+ \\ (\alpha c_a - \beta)^+ \end{pmatrix}$.

Figure 2.7 illustrates the field configuration described above.

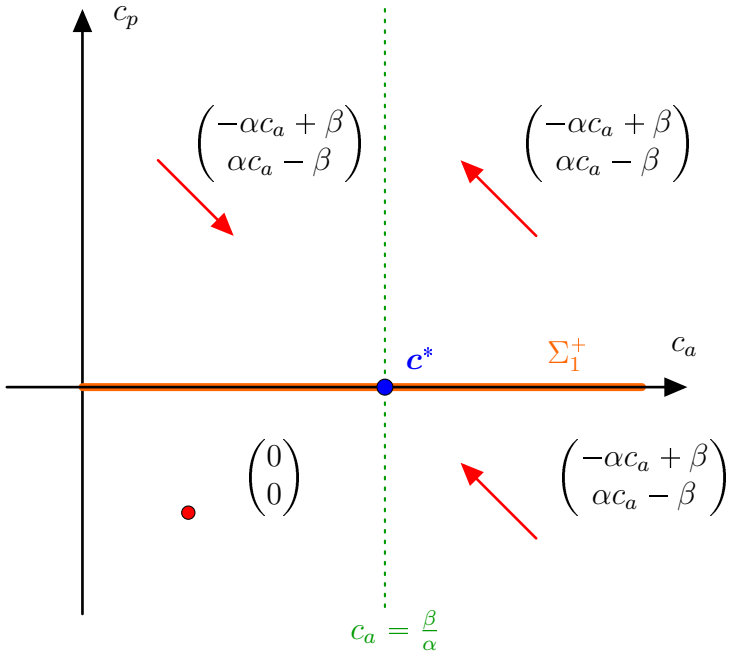


Figure 2.7 – Field configuration of the reference example in the tangent case

Let us consider the point $\boldsymbol{c}^* = (c_a^*, c_p^*) = (\frac{\beta}{\alpha}, 0)$. *The normal components* $f_2^1(\boldsymbol{c}^*)$ *and* $f_2^2(\boldsymbol{c}^*)$ *are equal to zero, so that the tangent conditions are satisfied. In this case the other components are also equal zero:,* $f_1^1(\boldsymbol{c}^*) = f_1^2(\boldsymbol{c}^*) = 0$.

Hence, $\boldsymbol{F}(\boldsymbol{c}^*) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$.

Uniqueness results for this particular example in the tangent case are obtained in [2] and [46].

In [2], the authors use the mass conservation to prove the one-sided Lipschitz continuity of the right-hand side, which yields the uniqueness. In fact, Theorem 2.3.1 is proved using two main arguments: one-sided Lipschitz property and a suitable change of variable allowing to consider the right orthogonal space. In view of [39], a future research direction would be to identify the right scalar product to prove the one-sided Lipschitz property of the right-hand side of (2.8) in a neighbourhood of a point in the tangent case for a general kinetic system.

Chapter 3

A projected dynamical system approach to mineral precipitation-dissolution reactions in geochemistry

This work corresponds to a preprint in view of a submission in a journal. The authors are: Jocelyne Erhel, Bastien Hamlat and Tangi Migot.

Abstract

In this paper, we study kinetically controlled precipitation-dissolution geochemical reactions. Ordinary differential equations are very often used to model aqueous reactions. Due to the nonsmoothness induced by the heterogeneous reactions, we propose a new model with a projected dynamical system. This new approach generalizes the existing methods that are limited to one mineral per reaction. We prove the existence of a solution, and we propose a fixed point projection-based method to solve the differential system discretized with an explicit scheme. Numerical computations on some examples illustrate this approach.

Introduction

Many environmental studies rely on modeling geochemical reactions such as CO_2 sequestration, water injection, and geothermal energy. In some of these applications, when the reaction rate is relatively slow compared to the environment, the reaction does not necessarily reach an equilibrium. In this case, we consider kinetically controlled chemical reactions. In several of the aforementioned applications, the chemical reactions are coupled with hydrodynamic processes leading for instance to a reactive transport problem [57, 17, 45, 36, 27]. It is of great interest to get an in-depth knowledge of the geochemical properties involved as it is very often the source of non-linearity or discontinuity in the system. In particular, this is true when considering heterogeneous reactions such as precipitation-dissolution reactions [65]. The difficulty here comes from the change of phase, which implies discontinuities in the reaction rate [40]. We also refer

the reader to the recent survey on computational methods for geochemistry and reactive transport modeling in [56] for more insights.

We focus here on kinetically-controlled reactions involving aqueous and mineral species. The concentration evolution of each species is governed by an ideal rate, depending on the law of mass action, while satisfying non-negativity and conservation constraints. As mentioned in [40], the resulting differential system might be discontinuous. This constitutes a real challenge both for the design of numerical methods and the study of the properties of the system. More sophisticated models can deal simultaneously with both kinetically controlled and equilibrium controlled reactions. We decided to focus on the former here for more simplicity. We refer to [45, 36, 27, 5] or recently to [35] for more details on this important subject.

Despite the difficulty of the problem, the past decade has seen some new developments in the study of such geochemical model [45, 12, 2, 46, 53]. These new developments are following up some early discussions in [40, 49, 70, 71, 72]. In the literature, the discontinuity induced by the phase change has been handled in several ways: an Heaviside function in [2, 53, 49, 70, 71, 72], a complementarity condition in [46], and an ordinary differential equation with discontinuous reaction rate in [40, 12]. In a recent article [46], the authors discuss the equivalence between these three approaches. In [45], based on a dynamical system with complementarity condition (in other words, a complementarity system), the authors study a semi-smooth Newton method to solve reactive transport problems with kinetic reactions.

A common point in the aforementioned articles is that the authors assume that there is at most one mineral per reaction. To the best of our knowledge, the only exception is [40], in which the authors focus on some examples. Our main motivation in this paper is to propose a new model that allows several minerals in the reactions.

Our new model is a projected dynamical system, which is a differential equation that simultaneously deals with the dynamics and the set of constraints. The difficulty here is that the right-hand side of the differential equation is discontinuous at the boundary of the set. To tackle this difficulty, we propose a numerical scheme based on an explicit discretization of the equation and a projection-type method. We prove the convergence of this numerical method, which proves the existence of a solution constructively. We show on several examples the results obtained by an implementation of this new approach. Moreover, we show through a classical example that our new model extends the complementarity system from [46].

The paper is organized as follows. In Section 1, we introduce the projected dynamical model for geochemistry and describe our assumptions. In Section 2, we study the connection between our new formulation and existing models used in the literature on two examples. In Section 3, we introduce a discretization of this projected dynamical system. We prove the convergence of the numerical scheme and, as a consequence, the existence of a solution of our model. In Section 4, we present an alternating projection algorithm to compute the discretized solution. In Section 5, we numerically solve several kinetic test cases. In particular, these examples illustrate that our model allows solving numerically kinetic systems with more than one mineral per reaction. Finally, we discuss some avenues for future work in Section 6.

3.1 A Projected Dynamical System Model

In this section, we first recall some characteristics of kinetic chemical systems. Then we give some constraints and introduce our new model.

3.1.1 Kinetic system

We denote by n the number of species taking part into m reactions, where we assume that $n > m$. The vector $x \in \mathbb{R}^n$ contains the species concentrations. We consider chemical reactions composed of aqueous components and minerals. There is no restriction on the number of minerals per reaction on either side of the reaction. We denote by $\mathcal{W} \subseteq \{1, \cdots, n\}$ the set of indices of the aqueous species.

$S \in \mathbb{Z}^{n\times m}$ denotes the stoichiometry matrix, where, by convention, the stoichiometric coefficients of the reactants are taken as negative integers. The sign convention is required here to define the rate of each reaction.

Let $Ker(S^T)$ be the kernel of S^T. A basis Q of $Ker(S^T)$ is a collection of vectors of so-called conservative quantities. Assuming that S is a full rank matrix, $Q \in \mathbb{R}^{n\times(n-m)}$ is also a full rank matrix satisfying

$$\mathrm{Im}(Q) = \mathrm{Ker}(S^T).$$

We assume that the activity of an aqueous species is equal to its concentration, whereas the activity of a mineral species is equal to 1. In these conditions, we use the kinetic mass action law to define the reaction rates $\tau(x) \in \mathbb{R}^m$. We get, for each $j = 1, \ldots, m$:

$$\tau_j(x) := k_{r,j} \prod_{\substack{i\in\mathcal{W} \\ s_{i,j}<0}} x_i^{-s_{i,j}} - k_{p,j} \prod_{\substack{i\in\mathcal{W} \\ s_{i,j}>0}} x_i^{s_{i,j}}, \tag{3.1}$$

where we denote by k_r (resp. k_p) the constant vector of the left side (resp. of the right side) of reactions corresponding to the reactants (resp. to the products).

The table below summarizes notations used throughout the paper.

Name	Definition
$n \in \mathbb{N}$	Number of species
$m \in \mathbb{N}$	Number of reactions
$S \in \mathbb{Z}^{n\times m}$	Stoichiometry matrix
$x \in \mathbb{R}^n$	Concentration

Name	Definition
$Q \in \mathbb{R}^{n\times(n-m)}$	Conserved quantities matrix
$k_p \in \mathbb{R}^m$	Products species constant
$k_r \in \mathbb{R}^m$	Reactants species constant
$\tau \in \mathbb{R}^m$	Reaction rate

3.1.2 Model

Reaction rates are driving the evolution of the system such that $\dot{x} = S\tau(x)$ and $Q^T\dot{x} = 0$ as long as $x > 0$. However, this evolution is perturbed (but not necessarily stopped) when a mineral species is vanishing, as illustrated for instance in [40, Exemple 3].

In order to design a model taking into account vanishing minerals, we first assume that the derivative of the concentration vector remains proportional to the reaction rate.

In other words, for each species $i \in \{1, \dots, n\}$, there exists $\lambda_i \in [0, 1]$ such that $\dot{x_i} = (S\tau(x))_i \lambda_i$. This can be rewritten as

$$\dot{x} \in D(x) := \{\text{diag}(S\tau(x))\lambda \text{ with } \lambda \in [0, 1]^n\}.^1$$

Moreover, concentrations must satisfy contraints. Starting from a non-negative initial condition $x_0 \in \mathbb{R}^n_+$, the vector $x(t)$ should remain non-negative for all $t \in \mathbb{R}_+$. Besides, conservative variables $Q^T x(t)$ should be invariant. Said otherwise, the concentration vector satisfies at any time

$$x(t) \in C_0 := \{x \in \mathbb{R}^n_+ \ : \ Q^T(x - x_0) = 0\}.$$

The following lemma links the conserved quantities to the derivative of x.

Lemma 3.1.1. *Let $x : [0, +\infty[\to \mathbb{R}^n$ be an absolutely continuous function such that $x(0) = x_0$. For almost every $t \in [0, +\infty[$, it holds true that*

$$Q^T(x(t) - x_0) = 0 \iff Q^T \dot{x}(t) = 0.$$

Proof. The function $y(t) := Q^T(x(t) - x_0)$ is also absolutely continuous with $\dot{y}(t) = Q^T \dot{x}(t)$ almost everywhere and $y(0) = 0$.

Then, using derivation properties, we get, almost everywhere, $y(t) = y_0 \iff \dot{y}(t) = 0$ and the desired result. □

Before introducing our model, we recall two classical definitions from convex analysis. The tangent cone to a convex set C at $x \in C$ is defined by

$$\mathcal{T}(C; x) = \{d : \exists d_\nu \to d \text{ and } \exists \tau_\nu \downarrow 0 \text{ with } x + \tau_\nu d_\nu \in C \ \forall \nu\}. \tag{3.2}$$

The projection of a vector v onto C is defined by

$$\text{Proj}(C; v) := \{w \in C : \|v - w\|_2 = \inf_{z \in C} \|v - z\|_2\}.$$

Clearly, C_0 is a non-empty convex set defined by linear inequality and equality constraints. Thus, the tangent cone to C_0 at x is non-empty, closed and convex and has the following explicit form:

$$\mathcal{T}(C_0; x) = \{d : Q^T d = 0, d_i \geqslant 0 \text{ for all } i \text{ such that } x_i = 0\}. \tag{3.3}$$

Hence $\mathcal{T}(C_0; x) \cap D(x)$ is a closed and convex set.

Now, we can define our model. Combining the two conditions described above, we get the following projected dynamical system:

$$\begin{cases} x(t) \in C_0, \\ \dot{x}(t) = \text{Proj}\left(D(x(t)) \cap \mathcal{T}(C_0, x(t)); S\tau(x(t))\right), \\ x(0) = x_0. \end{cases} \tag{3.4}$$

This model is similar to a classical projected dynamical system, with projection onto the tangent cone, which has a unique solution [61, 21]. However, here, projection is onto the intersection of $\mathcal{T}(C_0, x(t))$ with the set $D(x(t))$.

Therefore, we must prove the existence of a solution satisfying our projected dynamical system (3.4).

1. For every vector $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$, $\text{diag}(a)b = \text{diag}(b)a = (a_i b_i)_{i \in \{1, \cdots, n\}}$.

Definition 3.1.1. *Let $x_0 \in \mathbb{R}^n_+$. A function $x : [0, t_f] \to C_0$ is a solution of* (3.4) *if x is absolutely continuous and satisfies* (3.4) *for almost every $t \in [0, t_f]$.*

3.2 Link with Other Models

3.2.1 Link with a complementarity system

In the context of multiphase models, it is recurrent to use complementarity problems in the presence of heterogeneous reactions. In [46], the authors propose a complementarity formulation of kinetic systems, assuming one mineral per reaction and one reaction per mineral.

We focus on a reference example with an aqueous species W, a mineral species M, and one reaction:

$$W \rightleftharpoons M,$$

so that $S = (-1\ 1)^T$ and $Q = (1\ 1)^T$. The reaction rate is given by $\tau(x(t)) = k_r x_1(t) - k_p$, with two constants k_r and k_p.

The complementarity model introduced in [46] is written:

$$\begin{cases} \dot{x}_1(t) + \dot{x}_2(t) = 0, \\ 0 \leqslant x_2(t) \perp \dot{x}_2(t) - \tau(x(t)) \geqslant 0, \\ x(0) = x_0. \end{cases} \tag{3.5}$$

For this example, we prove that our model (3.4) is equivalent to the complementarity model (3.5).

Theorem 3.2.1. *The projected dynamical system* (3.4) *is pointwise equivalent to the complementarity system* (3.5).

Proof. Let $x : [0, t_f] \to \mathbb{R}^2_+$ be a Lipschitz continuous function. We can split $[0, t_f]$ into a finite number of subintervals such that, in each open interval I, x is differentiable and satisfies one of the three following conditions: either $\forall t \in I, x_1(t) = 0$ or $\forall t \in I, x_2(t) = 0$ or $\forall t \in I, x_1(t) > 0$ and $x_2(t) > 0$.

In the first case, since $x_1(t) = 0$ over I, it follows that $\forall t \in I, \dot{x}_1(t) = 0$. Hence, noticing that the constraint $\dot{x}_1(t) + \dot{x}_2(t) = 0$ is common to both problems, we get $\forall t \in I, \dot{x}_1(t) = \dot{x}_2(t) = 0$. Therefore, both problems are equivalent in this case. The second case is similar using x_2 instead of x_1.

In the third case, the complementarity problem (3.5) can be rewritten

$$\forall t \in I, \dot{x}(t) = (-\tau(x(t)), \tau(x(t)))^T.$$

On the other hand,

$$\mathrm{Proj}(D(x(t)) \cap \mathcal{T}(C_0; x(t)); (-\tau(x(t)), \tau(x(t)))^T) = (-\tau(x(t)), \tau(x(t)))^T,$$

so that the projected dynamical system is equivalent.

Finally, both models are equivalent in each open interval I, thus almost everywhere in $[0, t_f]$.

□

3.2.2 Link with a discontinuous model

In some kinetic models, positivity is ensured by limiting each reaction rate [12, 43]. In the model proposed in this paper, we choose to limit each time derivative.

For the reference example $W \rightleftharpoons M$ above, discontinuous models proposed in [12, 43] are equivalent to the complementarity problem [45]. So our model is also equivalent to discontinuous ones.

As said above, we also consider kinetic systems where reactions can involve several minerals and where a mineral can participate in several reactions. We consider here a kinetic system with two reactions involving three minerals [43]:

$$M_1 \rightleftharpoons M_2 \text{ and } M_2 \rightleftharpoons M_3.$$

The stoichiometry matrix is $S = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}^T$ and a basis of $Ker(S^T)$ is $Q = (1\ 1\ 1)^T$. The reaction rate is constant and given by $\tau(x) = k_r - k_p$. For this example, we choose k_r and k_p so that $\tau = (2\ 6)^T$ and the initial condition is $x_0 = (1\ 1\ 1)^T$. Then $S\tau = (-2\ \ -4\ 6)^T$.

The discontinuous model proposed in [43] has a unique solution, which is piecewise affine. It is plotted in Figure 3.1 and defined below:

$$x(t) = \begin{pmatrix} x_{*,1}(t) - 2\lambda_1(t)(t - z(t)) \\ x_{*,2}(t) - 4\lambda_2(t)(t - z(t)) \\ x_{*,3}(t) + 6\lambda_3(t)(t - z(t)) \end{pmatrix} \text{ with } \lambda(t) = \begin{cases} (1\ 1\ 1)^T, & 0 \leqslant t < 1/4, \\ (1\ 0\ \frac{1}{3})^T, & 1/4 \leqslant t < 1/2, \\ (0\ 0\ 0)^T, & 1/2 \leqslant t \leqslant t_f, \end{cases}$$

$$z(t) = \begin{cases} 0, & 0 \leqslant t < \frac{1}{4}, \\ \frac{1}{4}, & \frac{1}{4} \leqslant t < \frac{1}{2}, \\ \frac{1}{2}, & \frac{1}{2} \leqslant t \leqslant t_f, \end{cases} \quad \text{and} \quad x_*(t) = \begin{cases} (1\ 1\ 1)^T, & 0 \leqslant t < 1/4, \\ (\frac{1}{2}\ 0\ \frac{5}{2})^T, & 1/4 \leqslant t < 1/2, \\ (0\ 0\ 3)^T, & 1/2 \leqslant t \leqslant t_f. \end{cases}$$

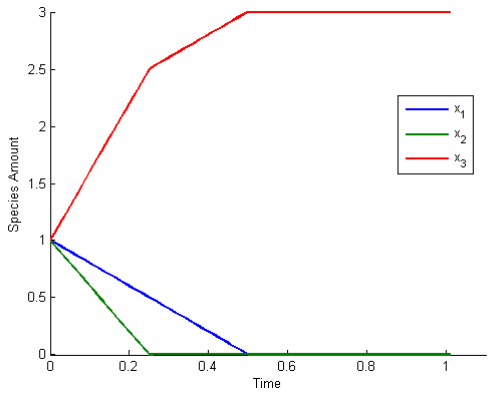


Figure 3.1 – Solution of the case $M_1M_2M_3$ considering $\tau = (2\ 6)^T$ and $x_0 = (1\ 1\ 1)^T$.

Now, we show that this function x is also solution of our model. It is clearly absolutely continuous and differentiable for all $t \in [0, +\infty[\setminus (\{\frac{1}{4}\} \cup \{\frac{1}{2}\})$.

The first conditions are easy to check: $\forall t, x(t) \in C_0$ and $\dot{x}(t) \in D(x(t)) \cap \mathcal{T}(C_0; x(t))$.

It remains to verify that $\dot{x}(t)$ is indeed the projection $y(t)$ of $S\tau(x(t))$ onto the set $D(x(t)) \cap \mathcal{T}(C_0; x(t))$ for all t.

This is clearly true for $t \in [0, \frac{1}{4}[$: $\dot{x}(t) = y(t)$.

Let us consider $t \in]\frac{1}{4}, \frac{1}{2}[$. Since $y(t) \in D(x(t))$ there exists $\mu \in [0,1]^3$ such that $y(t) = \text{diag}(S\tau)\mu = (-2\mu_1, -4\mu_2, 6\mu_3)^T$. Besides, since $\dot{x}(t) \in \mathcal{T}(C_0; x(t))$ and using (3.3), it follows that $\mu_2(t) = 0$ (recall that $x_2(\frac{1}{4}) = 0$ and $(S\tau(x(t)))_2 < 0$). Moreover, $Qy(t) = 0$ yields $\mu_1 = 3\mu_3$. Since $\mu \in [0,1]^3$, we get $\mu = (1\ 0\ \frac{1}{3})^T$. Thus $\dot{x}(t) = y(t)$.

Finally, for $t \in]\frac{1}{2}, 1]$, $D(x(t)) \cap \mathcal{T}(C_0; x(t))$ contains only $(0,0,0)^T$, since $x_1(\frac{1}{4}) = x_2(\frac{1}{4}) = 0$ and $(S\tau(x(t)))_1 < 0, (S\tau(x(t)))_2 < 0$.

In the numerical examples, we will also consider a similar system, with two aqueous species and one mineral:

$$W_1 \rightleftharpoons M \text{ and } M \rightleftharpoons W_2.$$

3.2.3 Link with a classical projected dynamical system

In this section, we illustrate by an example why the projection should contain explicitly the constraint $\dot{x}(t) \in D(x(t))$.

We introduce a classical projected dynamical system, which has a unique solution [61]:

$$\begin{cases} \dot{x}(t) = \text{Proj}\left(\mathcal{T}(C_0, x(t)); S\tau(x(t))\right), \\ x(0) = x_0. \end{cases} \tag{3.6}$$

We consider the same example as above:

$$M_1 \rightleftharpoons M_2 \text{ and } M_2 \rightleftharpoons M_3.$$

Figure 3.2 shows the solution of (3.6).

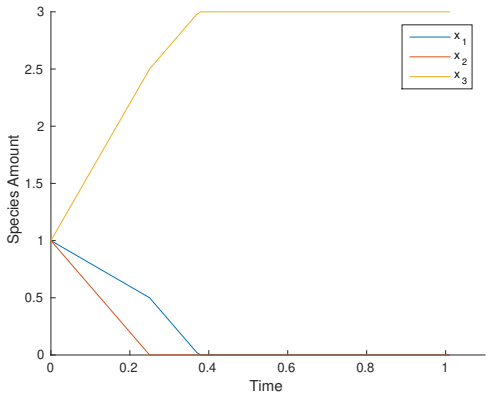


Figure 3.2 – Solution of the case $M_1M_2M_3$ considering $\tau = (2\ 6)^T$ and $x_0 = (1\ 1\ 1)^T$ without the constraint $\dot{x} \in D(x(t))$.

At time $t = \frac{1}{4}$, the second mineral M_2 is fully dissolved. Then $\dot{x}_1(t) = S\tau_1\lambda_1$, with $\lambda_1 > 1$ thus $\dot{x}(t) \notin D(x(t))$. In other words, model (3.6) does not satisfy the constraint for the time derivative in our model.

3.3 Existence of a Solution

In this section, we prove the existence of a solution to (3.4). We give a constructive proof by introducing a numerical scheme. For a given time step $h > 0$, we consider forward difference approximation of the derivative. Then, an approximate solution is defined by a projection and an affine interpolation. We prove that this explicit scheme converges as $h \to 0^+$ to a solution of (3.4).

Throughout this section, we denote by $\lfloor a \rfloor$ the largest integer smaller than a.

3.3.1 Explicit scheme and interpolation

Consider a fixed step size $h > 0$. For $k \in \mathbb{N}$ such that $kh \leqslant t_f$, let $\{x_k\}_k$ be the sequence defined such that x_0 is the initial point of (3.4) and

$$x_{k+1} = \mathrm{Proj}\left(K_k; x_k + hS\tau(x_k)\right), \tag{3.7}$$

where $K_k = C_0 \cap (x_k + hD(x_k))$. Our first observation is that this sequence is well-defined.

Proposition 3.3.1. *Let $h > 0$. For all $k = 0, \dots, \lfloor \frac{t_f}{h} \rfloor$, the projection in (3.7) exists and is unique.*

Proof. Note that for any k, the set K_k is non-empty as $x_k \in K_k$. Moreover, the set K_k is closed and convex. Therefore, the projection in (3.7) exists and is unique. □

We now sum up in the following lemma some properties of the sequence $\{x_k\}_k$ computed in (3.7).

Lemma 3.3.1. *Let $h > 0$, $k \in \mathbb{N}$ such that $hk \leqslant t_f$, and $x_k \in C_0$.*

Then

$$z \in K_k \iff \frac{z - x_k}{h} \in D(x_k) \cap \left\{ d : Q^T d = 0, d \geqslant -\frac{x_k}{h} \right\}.$$

Moreover, (3.7) is equivalent to

$$\begin{cases} x_{k+1} & = x_k + hd_k, \\ \frac{x_{k+1} - x_k}{h} & = \mathrm{Proj}\left(D(x_k) \cap \left\{ d : Q^T d = 0, d \geqslant -\frac{x_k}{h} \right\}; S\tau(x_k) \right). \end{cases} \tag{3.8}$$

Proof. Let z and d be related by $z = x_k + hd$. Since $x_k \in C_0$, it holds that

$$z \in C_0 \iff Q^T d = 0,\ d \geqslant -\frac{x_k}{h}$$

and

$$z \in x_k + hD(x_k) \iff d \in D(x_k).$$

Thus $z \in K_k$ if and only if $d \in D(x_k) \cap \{d : Q^T d = 0, d \geqslant -\frac{x_k}{h}\}$. Also, the projection in (3.7) can be rewritten as

$$\|x_{k+1} - x_k - hS\tau(x_k)\| \leqslant \|z - x_k - hS\tau(x_k)\|, \quad \forall z \in K_k.$$

Hence, dividing by $h > 0$ and defining d_k by $x_{k+1} = x_k + hd_k$, previous inequality is equivalent to

$$\|d_k - S\tau(x_k)\| \leqslant \|d - S\tau(x_k)\|, \quad \forall d \in D(x_k) \cap \{d : Q^T d = 0, d \geqslant -\frac{x_k}{h}\}.$$

The result follows again by definition of a projection. □

Using the sequence $\{x_k\}_k$, for $h > 0$ and $x_0 \in C_0$, we define an interpolation function $x_h : [0, t_f] \to \mathbb{R}^n$ by

$$x_h(t) = x_k + (x_{k+1} - x_k)\frac{t - kh}{h}, \text{ with } k = \lfloor \frac{t}{h} \rfloor. \tag{3.9}$$

Here, k is chosen such that $t \in [kh, (k+1)h[$.

The function x_h is piecewise affine and differentiable almost everywhere:

$$\dot{x}_h(t) = \frac{x_{k+1} - x_k}{h}, \text{ with } k = \lfloor \frac{t}{h} \rfloor, \text{ and } t \neq lh, l = 0, \dots, \lfloor \frac{t_f}{h} \rfloor. \tag{3.10}$$

The discrete function x_h satisfies an approximate model:

Lemma 3.3.2. *Let $h > 0$ and $x_0 \in C_0$. The discrete function x_h is solution of the following discrete projected dynamical system: for all $t \neq lh, l = 0, \dots, \lfloor \frac{t_f}{h} \rfloor$,*

$$\begin{cases} \dot{x_h}(t) = \text{Proj}\left(D(x_k) \cap \left\{d : Q^T d = 0, d \geqslant -\frac{x_k}{h}\right\}; S\tau(x_k)\right), \\ x_h(t) \in C_0, \\ x_h(0) = x_0, \end{cases} \tag{3.11}$$

where $k = \lfloor \frac{t}{h} \rfloor$.

Proof. For any $t \in [0, t_f]$, $x_h(t)$ is a convex combination of $x_k \in C_0$ and $x_{k+1} \in C_0$. Since C_0 is a convex set, $x_h(t) \in C_0$. By definition, $x_h(0) = x_0$.

Finally, by (3.10) and (3.8), the derivative satisfies the first equation of (3.11). □

3.3.2 Bounded constraints

In the special case where C_0 is bounded, a solution of the model is Lipschitz continuous. We introduce an assumption limiting the production of species. More precisely, we assume that the columns of Q^T are positively linearly independent.

Assumption 3.3.1.

$$Q^T d = 0, d \geqslant 0 \Rightarrow d = 0. \tag{3.12}$$

Using Motzkin alternative theorem [59], (3.12) is equivalent to the following condition:

$$\{d \mid Qd > 0\} \neq \varnothing.$$

Since Q is a basis of $Ker(S^T)$, it can also be written

$$\{q \in Ker(S^T), \mid q > 0\} \neq \varnothing.$$

This last condition is used for example in [46, Assumption 4.1]

The following lemma shows that, thanks to Assumption 3.3.1, the production of concentration is necessarily finite.

Lemma 3.3.3. *If Q satisfies* (3.12) *then C_0 is bounded.*

Proof. Remind that $C_0 = \{x : Q^T(x - x_0) = 0, x \geqslant 0\}$. Assume by contradiction that there exists a sequence $\{x_n\}$ with $x_n \in C_0$ for all $n \in \mathbb{N}$ such that $\lim_{n\to\infty} \|x_n\|_\infty = \infty$.

Let $d_n := \frac{x_n - x_0}{\|x_n\|_\infty}$. This sequence is bounded, therefore it converges, up to a subsequence, to some limit:

$$\lim_{n\to\infty} d_n = \lim_{n\to\infty} \frac{x_n}{\|x_n\|_\infty} = d \geqslant 0 \text{ with } d \neq 0.$$

We obtain $d \neq 0$ using that if $\lim_{n\to\infty} \|x_n\|_\infty = \infty$ then there exists a component of x_n such that, up to a subsequence, the limit of this component is infinite. So, there exists a non-zero component of d.

Besides, since $x_n \in C_0$ it holds true that $Q^T d_n = 0$ for all n and then $Q^T d = 0$. However, Assumption 3.3.1 gives that $d = 0$, which is a contradiction. □

The following example illustrates a system where production may arbitrarily grow and Assumption 3.3.1 does not hold.

Example 3.3.1. *Consider the following system of two reactions with three aqueous species:* $W_1 \rightleftharpoons W_2$ *and* $W_2 \rightleftharpoons 2W_1 + W_3$. *This system seems to create an infinite quantity of species, since whenever the first reaction consumes a quantity of* W_1, *then the second reaction doubles this quantity.*

In this case, the stoichiometry matrix and a basis of the kernel of its transpose are given by

$$S = \begin{pmatrix} -1 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \text{ and } Q = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Assumption 3.3.1 fails to hold here as $Q^T d = 0$ *for* $d = (1\ 1\ 2)^T > 0$.

A direct consequence of this lemma is that any element in the set $D(x)$ is also bounded under Assumption 3.3.1.

Lemma 3.3.4. *Assume that* (3.12) *holds true. There exists* $\rho > 0$ *such that, for any* $x \in C_0$ *and any* $d \in D(x)$, $\|d\| \leqslant \rho$.

Proof. Let $x \in C_0$ and $d \in D(x)$. By definition of D, there exists $\lambda \in [0, 1]^n$ such that $d = \mathrm{diag}(S\tau(x))\lambda$. Therefore, it holds

$$\|d\| = \|\mathrm{diag}(S\tau(x))\lambda\| \leqslant \|\lambda\|\|S\tau(x)\| \leqslant \|S\tau(x)\|.$$

Since C_0 is bounded by Lemma 3.3.3, and τ is continuous, then $\tau(x)$ is uniformly bounded. Hence, there exists $\rho > 0$ such that $\|d\| \leqslant \|S\tau(x)\| \leqslant \rho$. □

3.3.3 Convergence of the discrete function

Throughout this section, we make Assumption (3.3.1), such that C_0 is bounded. Thanks to this assumption, we will prove convergence of x_h towards a function x.

Proposition 3.3.2. *There exists a subsequence* $\mathcal{H}$ *and a Lipschitz continuous function* $x : [0, t_f] \to C_0 \subset \mathbb{R}^n$ *such that* $x_h \to_{\mathcal{H}} x$ *uniformly.*

Proof. By Lemma 3.3.3, the sequence $\{x_h\}_h$ is equi-bounded. Also, by Lemma 3.3.4, the sequence $\{x_h\}_h$ is equi-Lipschitz with constant ρ.

Then, all the assumptions of the Arzela-Ascoli theorem are satisfied. Therefore, there exists a subsequence $\mathcal{H}$, such that $h \to_{\mathcal{H}} 0$, and a Lipschitz mapping $x : [0, t_f] \to \mathbb{R}^n$ such that $x(0) = x_0$ and $\{x_h\}_{h\in\mathcal{H}}$ converges uniformly to x in $[0, t_f]$. Besides, since C_0 is closed, then $x(t) \in C_0$ for all $t \in [0, t_f]$. □

Since k depends on h and t, the sequence x_k converges.

Lemma 3.3.5. *Let $t \in [0, t_f]$. The sequence x_k with $k = \lfloor \frac{t}{h} \rfloor$ converges towards $x(t)$: $x_k \to_{\mathcal{H}} x(t)$.*

Proof. By definition of x_h in (3.9) and using Lemma 3.3.4, it follows

$$\|x_h(t) - x_k\| \leqslant \|(x_{k+1} - x_k)\frac{t - kh}{h}\| \leqslant \rho(t - kh) \leqslant \rho h.$$

Passing to the limit as $h \to 0^+$, and using convergence of $x_h(t)$ in Proposition 3.3.2, we get convergence of x_k towards $x(t)$. □

Another consequence of Proposition 3.3.2 is the weak convergence of the derivative.

Lemma 3.3.6. *$\dot{x}_h \to_{\mathcal{H}} \dot{x}$ weakly in $L^1([0, t_f], \mathbb{R}^n)$.*

Proof. The result is a consequence of a straightforward application of [47, Lemma 2.2]. □

Even though the convergence is weak, there exists a sequence of functions which converge strongly towards the derivative of x.

Lemma 3.3.7. *There exists a sequence of functions $\{y_h\}_h \to_{\mathcal{H}} \dot{x}$ and satisfying*

$$y_h = \sum_{l\in J_h} \alpha_l \dot{x}_{h_l}, \text{ with } \sum_{l\in J_h} \alpha_l = 1, \text{ and } \alpha_l \geqslant 0, \forall l \in J_h,$$

where J_h is a finite set of indices such that $(h_l)_{l\in J_h} \subset \mathcal{H}$.

Proof. The sequence $\{\dot{x}_h\}_{h\in\mathcal{H}}$ converges weakly to $\dot{x}$, by Lemma 3.3.6. Then, applying Mazur's Lemma, there exists a sequence $\{y_h\}_h$, with values in the convex hull of the set $\{\dot{x}_h\}_{h\in\mathcal{H}}$, which converges to $\dot{x}$. □

The following lemma will also be useful when passing to the limit in the constraint set. The set convergence has to be understood in the Painlevé-Kuratowski sense [64, Chapter 5.B].

Lemma 3.3.8. *Let $z \in C_0$ and $\{z_h\}_h \subset C_0$ be a sequence such that $z_h \to_{h\to 0^+} z$. Then*

$$\lim_{h\to 0^+} D(z_h) \cap \{d : Q^T d = 0, d \geqslant -\frac{z_h}{h}\} = D(z) \cap \{d : Q^T d = 0, d \geqslant -\lim_{h\to 0^+} \frac{z_h}{h}\}.$$

Proof. By Lemma 3.3.3 and Lemma 3.3.4, the sets C_0 and $D(z_h) \cap \{d : Q^T d = 0, d \geqslant -\frac{z_h}{h}\}$ are both non-empty and compact for all $h > 0$. Therefore, by [64, Theorem 5.7], it follows that $\limsup_{h\to 0^+} D(z_h) \cap \{d : Q^T d = 0, d \geqslant -\frac{z_h}{h}\} \subseteq D(z) \cap \{d : Q^T d = 0, d \geqslant -\lim_{h\to 0^+} \frac{z_h}{h}\}$.

On the other hand, let $\xi \in D(z) \cap \{d : Q^T d = 0, d \geqslant -\lim_{h\to 0^+} \frac{z_h}{h}\}$. Since $x \mapsto S\tau(x)$ is continuous, there exists a sequence $\xi_h \in D(z_h) \cap \{d : Q^T d = 0, d \geqslant -\frac{z_h}{h}\}$ which converges to ξ. Thus $D(z) \cap \{d : Q^T d = 0, d \geqslant -\lim_{h\to 0^+} \frac{z_h}{h}\} \subseteq \liminf_{h\to 0^+} D(z_h) \cap \{d : Q^T d = 0, d \geqslant -\frac{z_h}{h}\}$.

All in all, the limit exists and the result follows. □

3.3.4 Existence of a solution

Again, we make Assumption 3.3.1 throughout this section. Denote by $I := \{t \in [0, t_f] : \dot{x}(t) \text{ exists}\}$. By Sard's theorem, the set $[0, t_f]\backslash I$ has a zero measure.

We will prove that the limit x of x_h is a solution of the model. First, we show that it satisfies the constraints.

Lemma 3.3.9. *For all $t \in I$, $\dot{x}(t) \in D(x(t)) \cap \{d : Q^T d = 0, d \geqslant -\lim_{\mathcal{H}} \frac{x_{k_l}}{h_l}\}$, with $l \in J_h$ and $k_l = \lfloor \frac{t}{h_l} \rfloor$.*

Proof. Let $t \in I$ and $h > 0$. Considering y_h defined in Lemma 3.3.7, and using (3.11), we get

$$y_h(t) \in \operatorname{conv}\left(\bigcup_{l \in J_h} D(x_{k_l}) \cap \{d : Q^T d = 0, d \geqslant -\frac{x_{k_l}}{h_l}\} \right),$$

where conv denotes the convex hull and $k_l = \lfloor \frac{t}{h_l} \rfloor$.

We note that $h_l \to_{\mathcal{H}} 0$ and, using Lemma $x_{k_l} \to_{\mathcal{H}} x(t)$. Thus the sequence $x_{k_l}(t)$ satisfies assumptions of Lemma 3.3.8. Passing to the limit in y_h, we get the result. □

Lemma 3.3.10. *For all $t \in I$, $\dot{x}(t) \in \mathcal{T}(C_0; x(t))$.*

Proof. Recall that the tangent cone to C_0 is given by (3.3). By Lemma 3.3.9, $Q^T \dot{x}(t) = 0$.

By Lemma 3.3.9, $\dot{x}(t) \in D(x(t))$ so that $\dot{x}_i(t) = (S\tau(x(t)))_i \lambda_i$ for $i = 1, \ldots, n$, with $0 \leqslant \lambda_i \leqslant 1$. Let i such that $x_i(t) = 0$. We must show that $\dot{x}_i(t) \geqslant 0$.

If $(S\tau(x(t)))_i \geqslant 0$, then $\dot{x}_i(t) \geqslant 0$. On the other hand, if $(S\tau(x(t)))_i < 0$, by continuity of τ and x, there is an open interval U centered in t such that $(S\tau(x(s)))_i < 0$ so that $\dot{x}_i(s) \leqslant 0$ for all $s \in U \cap I$. Also, $x_i(t) = 0$ and $x_i(s) \geqslant 0$. Since x is absolutely continuous, we get $\dot{x}_i(t) = 0$. This concludes the proof. □

Lemma 3.3.11. *The limit x satisfies the inclusion below, for $t \in I$ and $l \in J_h$:*

$$D(x(t)) \cap \mathcal{T}(C_0; x(t)) \subseteq D(x(t)) \cap \{d : Q^T d = 0, d \geqslant -\lim_{h \in \mathcal{H}} \frac{x_{k_l}}{h_l}\}.$$

Proof. Since $x_{k_l} \geqslant 0$, the limit satisfies $-\lim_{h \in \mathcal{H}} \frac{x_{k_l}}{h_l} \leqslant 0$. Let i such that $x_i(t) > 0$, then $-\lim_{h \in \mathcal{H}} \frac{(x_{k_l})_i}{h_l} = -\infty$.

So, if $d \in \mathcal{T}(C_0; x(t))$, it satisfies $d \geqslant -\lim_{h \in \mathcal{H}} \frac{x_{k_l}}{h_l}$. □

We are now able to prove the existence of a solution to our model.

Theorem 3.3.1. *The limit function x is a solution of* (3.4).

Proof. The function $x : [0, t_f] \to C_0$ is a Lipschitz continuous function such that $x(0) = x_0$. It satisfies the constraints of the model, so that it remains to verify the projection property.

Considering y_h defined in Lemma 3.3.7, and using the triangle inequality, we get

$$\|S\tau(x_k) - y_h(t)\| \leqslant \max_{l \in J_h} \|S\tau(x_k) - \dot{x}_{h_l}(t)\|. \tag{3.13}$$

Let $j \in J_h$ which maximizes this inequality and $d \in D(x_{k_j}) \cap \{d : Q^T d = 0, d \geqslant -\frac{x_{k_j}}{h_j}\}$. Since x_{h_j} is solution of (3.11), it satisfies

$$\|S\tau(x_{k_j}) - \dot{x}_{h_j}(t)\| \leqslant \|S\tau(x_{k_j}) - d\|,$$

therefore

$$\|S\tau(x_k) - y_h(t)\| \leqslant \|S\tau(x_k) - S\tau(x_{k_j})\| + \|S\tau(x_{k_j}) - d\|.$$

Thanks to Lemma 3.3.5, $\lim_{h\in\mathcal{H}} x_{k_j} = \lim_{h\in\mathcal{H}} x_k = x(t)$. Using Lemma 3.3.7 and Lemma 3.3.8, we can pass to the limit:

$$\|S\tau(x(t)) - \dot{x}(t)\| \ \leqslant \ \|S\tau(x(t)) - d\|, d \in D(x(t)) \cap \{d : Q^T d = 0, d \geqslant -\lim_{h\in\mathcal{H}} \frac{x_{k_j}}{h_j}\}.$$

Finally, Lemma 3.3.11 provides the result. □

3.4 An Alternating Projection Algorithm

We conclude this analysis by presenting an alternating projection method to compute the sequence x_k. This iterative algorithm consists in computing a sequence $\{y_j\}_j$, starting from $x_k + hS\tau(x_k)$, which converges towards x_{k+1}. The sequence is defined as follows:

$$\begin{aligned}
&y_0 = x_k + hS\tau(x_k), e_0 = e_{-1} = 0,\\
&y_{2j+1} = \mathrm{Proj}(C_0; y_{2j} + e_{2j-1}),\\
&e_{2j+1} = y_{2j} + e_{2j-1} - y_{2j+1},\\
&y_{2j+2} = x_k + h\mathrm{Proj}\left(D(x_k); \frac{y_{2j+1} + e_{2j} - x_k}{h}\right),\\
&e_{2j+2} = y_{2j+1} + e_{2j} - y_{2j+2}.
\end{aligned} \tag{3.14}$$

The following result proves the convergence of the sequence (3.14).

Theorem 3.4.1. *The sequence $\{y_j\}_j$ computed by* (3.14) *converges to x_{k+1} satisfying* (3.7).

Proof. Using [28, Proposition 2.7 (ii) and (iv)], the fourth line of (3.14) can be rewritten

$$y_{2j+2} = \mathrm{Proj}(x_k + hD(x_k); y_{2j+1} + e_{2j}).$$

The algorithm is a direct application of the Boyle-Dykstra Theorem [28, Theorem 9.24]. We already noted that C_0 and $x_k + hD(x_k)$ are closed convex set and that their intersection K_k is non-empty. Therefore, the sequence y_j converges and the limit is nothing else than x_{k+1}. □

3.5 Numerical Experiments

We developed a software in Matlab in order to compute the discrete solution given by (3.9). For each $k = 0, \ldots, \lfloor \frac{t_f}{h} \rfloor$, and given x_k, the new iterate x_{k+1} is computed with

the alternating projection algorithm (3.14), implemented with an error tolerance of $\epsilon > 0$ and a maximal number of iterations j_{max} as in Algorithm 1.

Data: starting vector x_k; iteration threshold ϵ;
Begin ;
Set $j := 0$, $y_0 := x_k + hS\tau(x_k)$, $e_0 := e_{-1} := 0$;
repeat
  $y_{2j+1} = \text{Proj}(C_0; y_{2j} + e_{2j-1})$;
  $e_{2j+1} = y_{2j} + e_{2j-1} - y_{2j+1}$;
  $d_j := \text{Proj}\left(D(x_k); \dfrac{y_{2j+1} + e_{2j} - x_k}{h}\right)$;
  $y_{2j+2} := x_k + h\, d_j$;
  $e_{2j+2} = y_{2j+1} + e_{2j} - y_{2j+2}$;
  $j := j + 1$;
until $\|y_{2j} - y_{2j-1}\| \leqslant \epsilon$;
return: $x_{k+1} := y_{2j}$.

Algorithm 1: Alternating projection algorithm to (3.7).

Projections are computed using Matlab *lsqlin* function. In all experiments, the threshold ϵ is set to 10^{-7}, while the step size h is 10^{-2} in the time interval $[0, 1]$.

The sequel is divided into five examples illustrating the new approach discussed in this article. In all these examples, Algorithm 1 converges to a solution.

3.5.1 Example W_1W_2

The first example is a single reaction with two aqueous species:

$$W_1 \rightleftharpoons W_2.$$

The stoichiometric matrix is $S = (-1\ 1)^T$, and a basis of $\ker(S^T)$ is $Q = (1\ 1)^T$. The reaction rate is $\tau(x) = k_r x_1 - k_p x_2$.

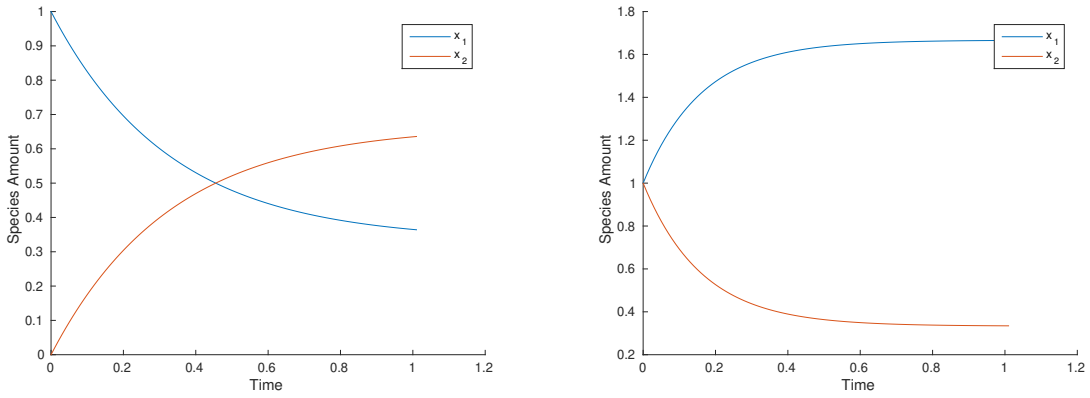


Figure 3.3 – Example W_1W_2. On the left: $k_r = 2, k_p = 1$ and $x_0 = (1\ 0)^T$; on the right: $k_r = 1, k_p = 5$ and $x_0 = (1\ 1)^T$.

Figure 3.3 shows the results on two different instances. In the first case (left), the second species W_2 is initially absent, then increases as the first species W_1 decreases. The

second case (right) illustrates an increase of W_1 with a decrease of W_2. In both cases, the quantities tend to an equilibrium at infinite time. Asymptotic values are given by $\tau(x) = 0$ and $Q^T x = Q^T x_0$.

3.5.2 Example WM

The following classical example, considered earlier in Section 3.2.1, is a single reaction with one aqueous species W and one mineral M:

$$W \rightleftharpoons M.$$

As in the previous test case, $S = (-1\ 1)^T$, and, $Q = (1\ 1)^T$. The reaction rate is $\tau(x) = k_r x_1 - k_p$.

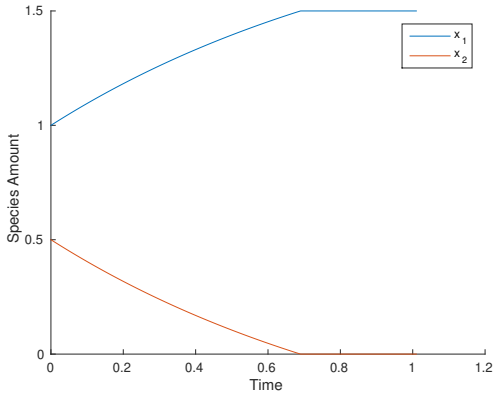


Figure 3.4 – Example WM, with $k_r = 1, k_p = 2$ and $x_0 = (1\ 0.5)^T$.

Figure 3.4 illustrates dissolution of the mineral M. As expected, the chemical reaction stops when M is fully dissolved. Asymptotic equilibrium is given by $x_2 = 0$ and $Q^T x = Q^T x_0$.

3.5.3 Example $M_1M_2M_3$

The following example is now two reactions involving three mineral:

$$M_1 \rightleftharpoons M_2 \text{ and } M_2 \rightleftharpoons M_3.$$

The stoichiometric matrix is $S = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}^T$, and, $Q = (1\ 1\ 1)^T$ is a non-zero element in the kernel of S^T. The reaction rate is $\tau(x) = k_r - k_p$. This exemple is used in Section 3.2.2 to illustrate a solution of (3.4) with discontinuous reaction rates. Figure 3.5 shows the evolution of the three solutions. In this more sophisticated example, we see a three stages reaction. Each stage is affected by the disappearance of one mineral and then follow a different reaction rate. The mineral M_2, which appears in both reactions vanishes first. Then, the mineral M_1 reduces thanks to the constraint on the conservation until only the mineral M_3 remains and the reaction stops.

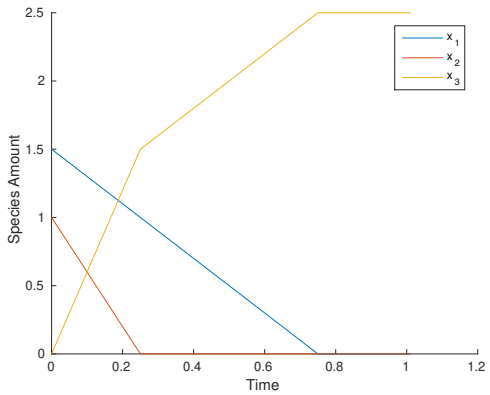


Figure 3.5 – Example $M_1M_2M_3$. This example considers $k_r = (3\ 7)^T, k_p = (1\ 1)^T$ and $x_0 = (1.5\ 1\ 0)^T$.

3.5.4 Example WM_1M_2

The following example considers again two reactions, but involving two minerals and one aqueous species:

$$W \rightleftharpoons M_1 \text{ and } M_1 \rightleftharpoons M_2.$$

The stoichiometric matrix is $S = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}^T$, and, $Q = (1\ 1\ 1)^T$ is a non-zero element in the kernel of S^T. The reaction rate is $\tau(x) = (k_{r,1}x_1 - k_{p,1}\ k_{r,2} - k_{p,2})^T$.

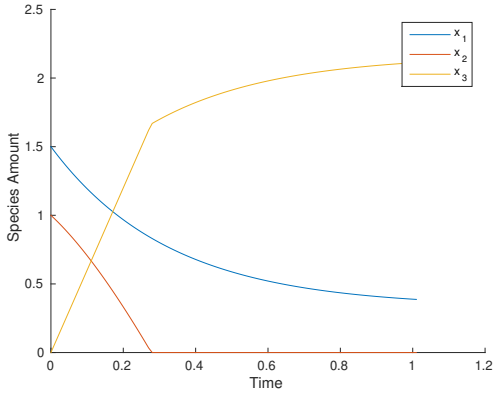


Figure 3.6 – Example WM_1M_2, with $k_r = (3\ 7)^T, k_p = (1\ 1)^T$ and $x_0 = (1.5\ 1\ 0)^T$.

Figure 3.6 shows the dynamics of the chemical system. The aqueous species W decreases and produces the first mineral M_1 which is dissolved into the second mineral M_2. Under the conditions chosen, M_2 is fully dissolved in a finite time. However, because the first reaction continues, the second reaction can also continue, at the reaction rate τ_1 instead of τ_2. In some sense, the system behaves as if W precipitates into mineral M_2, at rate τ_1. Asymptotic equilibrium, at infinite time, is given by $x_2 = 0$, $Q^T x = Q^T x_0$ and $\tau_1(x_1) = 0$.

3.6 Concluding Discussion

The study presented in this article introduces a new model for precipitation-dissolution reactions in the form of a projected dynamical system. The main benefit of this new approach is to offer the possibility to consider reactions with several minerals. This new model extends the complementarity system model for one mineral introduced in [46]. We presented a numerical scheme to find a solution to the differential system and proved its convergence. The existence of solutions of the equation follows as a consequence of the convergence. The numerical approach and the model have been validated on some examples.

Further researches may focus on the numerical perspectives offered by this work. The actual proposed method is based on an explicit discretization. The perspective to use an implicit scheme is an interesting path. Although in this case, the projection on D becomes more difficult.

The motivation to consider kinetic reactions is to assume that the reaction is too slow to attain its equilibrium. However, assuming that after a long time, the system reaches a stable state, we should obtain a connection with models for equilibrium reactions [35].

Finally, an essential motivation in this work is to apply this new technique in reactive transport modeling.

Conclusion et perspectives

Résumé des contributions de la thèse

Dans cette thèse nous apportons des réponses pour la modélisation des réactions de cinétiques chimiques multiphasiques. Nous proposons plusieurs types de modèle de limitation des vitesses de réaction cinétique et plusieurs manières de limiter la vitesse des concentrations des espèces chimiques. Dans le chapitre 1, nous proposons une approche par régularisation d'un modèle de surface réactive discontinue amont et un modèle limite de type inclusion différentielle. Dans le chapitre 2, nous proposons une approche discontinue d'un modèle de surface réactive amont, en utilisant la théorie de Filippov. Ces formulations sont basées sur une limitation amont des vitesses de réaction. Dans le chapitre 3, nous proposons une nouvelle approche de limitation en modulant directement la vitesse de chaque espèce à l'aide d'une dynamique projetée. Ce troisième modèle est différent des deux premiers car la limitation est appliquée espèce par espèce et non réaction par réaction, la stœchiométrie globale étant maintenue par l'ajout de contraintes de conservation.

Pour chacun des trois modèles, nous avons prouvé des résultats d'existence et de positivité des solutions. De plus, dans le chapitre 2, nous avons montré l'unicité des solutions du modèle développé pour le cas où un seul minéral disparaît et sous l'hypothèse que le champ ne soit pas tangent. Nous montrons que le modèle développé dans le chapitre 2 est une version plus précise du modèle développé dans le chapitre 1. Dans le chapitre 3, nous montrons sur des exemples choisis les différences et liens entre le modèle développé et d'autres modèles de cinétiques (modèle de type complémentarité, modèle discontinue du chapitre 1 et modèle projeté sans la contrainte sur la vitesse des réactions).

Enfin, nous proposons deux méthodes numériques de résolution des cinétiques : une basée sur le modèle régularisé du chapitre 1 et l'autre sur le modèle avec dynamique projeté du chapitre 3. Dans le chapitre 1, nous avons prouvé la convergence des solutions du modèle régularisé vers une solution d'un modèle de type inclusion différentielle. Dans le chapitre 3, nous avons montré la convergence du schéma numérique basé sur un algorithme à double projections. Nous illustrons, pour les deux méthodes, les résultats obtenus sur des cas tests unitaires variés issus de la cinétique chimique.

Travaux en cours et perspectives

Les trois premières perspectives font l'objet de travaux en cours.

Une première perspective est l'étude de la formulation de type inclusion différentielle

du chapitre 2 dans le cas où plusieurs minéraux disparaissent. Une étude préliminaire de classification des cas a été réalisée (configurations des champs possibles). Une étude du cas attractif en codimension 2 montre que le convexifié de Filippov ne détermine pas une valeur unique du second membre. Une sélection du second membre s'impose alors pour espérer l'unicité de la solution. Il existe plusieurs méthodes de sélection de $\boldsymbol{\lambda}$ dont l'une des plus connues est le "sigmoid blending" [4] and [32]. En chimie, la nature même du problème nous restreint au domaine des concentrations positives; en effet nous avons montré dans les chapitre 1 et 2 que le champs sont essentiellement positifs. Nous pouvons donc proposer une nouvelle sélection récursive du second membre adaptée aux problèmes de cinétique chimique comme combinaison convexe des champs de même codimension dans le domaine positif et du champ dans le domaine négatif (le champ du domaine pour lequel tous les minéraux ont une concentration négative). Cependant, une difficulté réside dans la démonstration de l'unicité de la solution après sélection du second membre. Une piste intéressante serait alors l'utilisation du théorème général d'unicité de Filippov [39, Theorem 4 Chapter 2 Section 10]. Plus généralement, on peut se poser la question pour un nombre quelconque de minéraux. On peut définir une version plus générale des conditions d'attractivité, mieux adaptée aux problèmes de cinétiques (voir chapitre 1 pour le cas avec un minéral), et démontrer une version générale du Théorème 2.3.2 (sliding mode).

Une deuxième perspective est d'établir le lien entre la formulation régularisée du chapitre 1 (selon le choix de la fonction de mélange pour $\boldsymbol{u_\epsilon}$) et la formulation au sens de Filippov du chapitre 2 (selon le choix du parametre du convexifié $\boldsymbol{\lambda}$). Dans le chapitre 1, nous avons montré la convergence faible-* de $\boldsymbol{u_\epsilon}$ vers $\boldsymbol{u}$. Une première difficulté est d'expliciter la formule de $\boldsymbol{u}$ en fonction du choix de la fonction de mélange pour $\boldsymbol{u_\epsilon}$. Des premiers travaux numériques nous permettent de conjecturer que si $u_{\epsilon,j} = \min_{i \in \mathcal{M}_j^-}(H_\epsilon(c_{\epsilon,i}))$ alors $u_j = \min_{i \in \mathcal{M}_j^-}(H(c_i))$ (de même pour le produit), avec H la fonction de Heaviside indéterminée en 0. Une idée serait d'effectuer une analyse de convergence forte afin de prouver cette conjecture. En supposant la conjecture vraie, on peut montrer que si le second membre limite du chapitre 1, $f(u, v, c)$, et le second membre du chapitre 2, $f(\lambda, c)$, sont égaux alors il existe une application linéaire inversible (dépendant seulement du nombre de minéraux N_p) entre $\tilde{H}$ et $\boldsymbol{\lambda}$, i.e. $\tilde{H} = A_{2^{N_p}}\lambda$ où $\tilde{H}$ est défini par

$$\tilde{H} = \big(1, H(c_1), \cdots, H(c_{N_p}), \min(H(c_1), H(c_1)), \cdots, \min(H(c_{N_p-1}), H(c_{N_p})), \cdots, \min(H(c_1), \cdots, H(c_{N_p}))\big)$$

La propriété est également vraie dans le cas où $\tilde{H}$ est défini avec la fonction produit. Plus généralement, cette relation nous permet de calculer des valeurs de $\tilde{H}$ connaissant λ. En particulier, on retrouve le λ sigmoid blending si on fait le choix produit pour le mélange. De plus, la détermination des valeurs de λ nous permet d'obtenir les bonnes valeurs en zéro des fonctions de Heaviside $H(c_i)$, $i = 1, \cdots, N_p$.

Une troisième perspective serait de proposer une formulation de type complémentarité générale. Dans [46], les auteurs proposent une formulation de type complémentarité pour des réactions contenant un unique minéral. Nous avons mené des premiers travaux et obtenu une formulation pour des réactions contenant un nombre quelconque de minéraux

:

pour une réaction de cinétique générale,

$$\sum_{i\in\mathcal{A}^-} |s_i|A_i + \sum_{k\in\mathcal{M}^-} |s_k|M_k \rightleftarrows \sum_{i\in\mathcal{A}^+} |s_i|A_i + \sum_{k\in\mathcal{M}^+} |s_k|M_k.$$

Soit $1 \in \mathcal{M}^-$, $Eb_i = \frac{\mathrm{d}c_i}{\mathrm{d}t} + |s_i|\tau$, $\forall i \in \mathcal{A}^- \cup \mathcal{M}^-$ et $Eb_i = \frac{\mathrm{d}c_i}{\mathrm{d}t} - |s_i|\tau$, $\forall i \in \mathcal{A}^+ \cup \mathcal{M}^+$.

$$\left\{\begin{array}{lcll} Eb_1 - Eb_i & = & 0, & \forall i \in \mathcal{A}^- \cup \mathcal{M}^- \backslash \{1\} \\ Eb_1 + Eb_i & = & 0, & \forall i \in \mathcal{A}^+ \cup \mathcal{M}^+ \\ med(\min\limits_{i\in\mathcal{M}^-}(c_i)\,,\, -\min\limits_{i\in\mathcal{M}^+}(c_i)\,, Eb_1\,) & = & 0, & \end{array}\right. \tag{3.15}$$

avec *med* la fonction qui renvoie la valeur médiane des trois quantitées (la deuxième valeur une fois ordonné).

Cependant, considérer un nombre quelconque de réactions n'est pas direct. Une piste serait d'étudier les liens entre les formulations de types EDOs discontinues, complémentarité et dynamique projetée, afin d'obtenir une formulation de type complémentarité équivalente aux modèles du chapitre 1 ou 3. Une autre piste serait de travailler directement sur les contraintes du système.

Une quatrième perspective serait de proposer un modèle général de transport réactif couplant équilibre et cinétique chimique dans l'objectif d'effectuer de la simulation numérique. Nous avons implémenté la loi de cinétique régularisée du chapitre 1 dans le logiciel de transport réactif IFPEN Geoxim combinée avec une formulation multiphasique de type Coats existante. Pour cela, nous avons dû adapter notre modèle à la structure du logiciel et réaliser le couplage avec la partie déjà existante, transport réactif avec équilibres chimiques. Cela constitue une première implémentation globalement implicite d'un modèle de transport réactif cinétique avec un nombre de minéraux quelconque par réaction. Des premières simulations numériques sur un maillage 2D ont été effectués pour valider l'approche sur des cas unitaires présentés lors du congrès CMWR 2018 (Computational Methods in Water Resources). Un premier travail envisagé porterait sur l'amélioration et la validation de la méthode numérique développée. Un objectif supplémentaire serait de considérer des problèmes issus de la physique ou de l'industrie. Pour cela, une première étape serait de valider et d'adapter les méthodes de résolution numérique pour les cinétiques chimiques (voir chapitre 1 et 3) à des problématiques appliquées en géosciences ou dans d'autres domaines avec des processus et des données réalistes. Un second travail envisagé porterait sur une analyse théorique des propriétés du système de transport-chimie couplé, en discret en en continu.

Bibliography

[1] V. Acary and B. Brogliato, Numerical Methods for Nonsmooth Dynamical Systems: Applications in Mechanics and Electronics, 2008.

[2] A. Agosti, L. Formaggia, and A. Scotti, Analysis of a model for precipitation and dissolution coupled with a darcy flux, Journal of Mathematical Analysis and Applications, 431 (2015), pp. 752 – 781.

[3] J. Alexander and T. I. Seidman, Sliding modes in intersecting switching surfaces, Houston journal of mathematics, (1995).

[4] ——, Sliding modes in intersecting switching surfaces, i: Blending, Houston journal of mathematics, (1998).

[5] L. Amir and M. Kern, A global method for coupling transport with chemistry in heterogeneous porous media, Computational Geosciences, 14 (2010), pp. 465–481.

[6] D. Angeli, A tutorial on chemical reaction network dynamics, European Journal of Control - EUR J CONTROL, 15 (2009), pp. 398–406.

[7] J. P. Aubin and A. Cellina, Differential Inclusions, Springer, Berlin, Heidelberg, 1984.

[8] H. Bauschke, J. Borwein, and A. S. Lewis, On the method of cyclic projections for convex sets in hilbert space, 1994.

[9] H. H. Bauschke and J. M. Borwein, On the convergence of von neumann's alternating projection algorithm for two sets, Set-Valued Analysis, 1 (1993), pp. 185–212.

[10] C. M. Bethke, Geochemical reaction modeling, 1996.

[11] O. Bildstein, Modelisation geochimique des interactions eau-gaz-roche application a la diagenese minerale dans les reservoirs geologiques, PhD thesis, 1998. Thèse de doctorat dirigée par Fritz, Bertrand Géochimie Strasbourg 1 1998.

[12] N. Bouillard, R. Eymard, R. Herbin, and P. Montarnal, Diffusion with dissolution and precipitation in a porous medium: Mathematical analysis and numerical approximation of a simplified model, ESAIM: Mathematical Modelling and Numerical Analysis, 41 (2007), pp. 975–1000.

[13] B. Brogliato, A. Daniilidis, C. Lemaréchal, and V. Acary, On the equivalence between complementarity systems, projected systems and differential inclusions, Systems & Control Letters, 55 (2006), pp. 45–51.

[14] F. Brunner, J. Fischer, and P. Knabner, Analysis of a modified second-order mixed hybrid bdm finite element method for transport problems in divergence form, SIAM Journal on Numerical Analysis, 54 (2016), p. 2359–2378.

[15] H. Buchholzer, C. Kanzow, P. Knabner, and S. Kräutle, The semismooth Newton method for the solution of reactive transport problems including mineral precipitation-dissolution reactions, Computational Optimization and Applications, 50 (2011), pp. 193–221.

[16] J. Carrayrou, Modélisation des phénomènes de transport réactif en milieu poreux saturé, HDR Thesis, Université de Strasbourg, 2018.

[17] J. Carrayrou, J. Hoffmann, P. Knabner, S. Kräutle, C. de Dieuleveult, J. Erhel, J. Van der Lee, V. Lagneau, K. U. Mayer, and K. T. B. MacQuarrie, Comparison of numerical methods for simulating strongly nonlinear and heterogeneous reactive transport problems–the MoMas benchmark case, Computational Geosciences, 14 (2010), pp. 483–502.

[18] V. Chellaboina, S. P. Bhat, W. M. Haddad, and D. S. Bernstein, Modeling and analysis of mass-action kinetics, IEEE Control Systems Magazine, (2009).

[19] K. H. Coats, An equation of state compositional model, (1980).

[20] M.-G. Cojocaru, Infinite-dimensional projected dynamics and the 1-dimensional obstacle problem, Journal of Function Spaces, 3 (2005), pp. 251–262.

[21] M. G. Cojocaru, P. Daniele, and A. Nagurney, Projected dynamical systems and evolutionary variational inequalities via Hilbert spaces with applications, Journal of Optimization Theory and Applications, 127 (2005), pp. 549–563.

[22] M.-G. Cojocaru and L. Jonker, Existence of solutions to projected differential equations in hilbert spaces, Proceedings of the American Mathematical Society, 132 (2004), pp. 183–193.

[23] J. Corvisier, Implémentation des phénomènes de germination / mûrissement / croissance des phases solides secondaires dans un modèle de transport-réactif en milieu poreux géologique. Développement du code de calcul Archimède, PhD thesis, 2006. Thèse de doctorat dirigée par Guy, Bernard et Garcia, Daniel Sciences de la Terre Saint-Etienne, EMSE 2006.

[24] H. Darcy, Les fontaines publiques de la ville de Dijon, Victor Dalmont, Paris, 1856.

[25] C. de Dieuleveult, Un modèle numérique global et performant pour le couplage géochimie-transport, PhD thesis, 2008. Thèse de doctorat dirigée par Erhel, Jocelyne Informatique Rennes 1 2008.

[26] C. de Dieuleveult and J. Erhel, A global approach to reactive transport: application to the momas benchmark, Computational Geosciences, 14 (2010), pp. 451–464.

[27] C. de Dieuleveult, J. Erhel, and M. Kern, A global strategy for solving reactive transport equations, Journal of Computational Physics, 228 (2009), pp. 6395–6410.

[28] F. R. Deutsch, Best approximation in inner product spaces, Springer Science & Business Media, 2012.

[29] L. Dieci, C. Elia, and L. Lopez, A Filippov sliding vector field on an attracting co-dimension 2 discontinuity surface, and a limited loss-of-attractivity analysis, Journal of Differential Equations, (2013).

[30] L. Dieci and N. Guglielmi, Regularizing piecewise smooth differential systems: Co-dimension 2 discontinuity surface, Springer, (2013).

[31] L. Dieci and L. Lopez, Sliding motion in Filippov differential systems: Theoretical results and a computational approach, SIAM Journal on Numerical Analysis, 47 (2009), pp. 2023–2051.

[32] ——, Sliding motion on discontinuity surfaces of high co-dimension. a construction for selecting a Filippov vector field, Numerische Mathematik, 117 (2011), pp. 779–811.

[33] ——, A survey of numerical methods for ivps of odes with discontinuous right-hand side, Journal of Computational and Applied Mathematics, 236 (2012), pp. 3967 – 3991.

[34] J. Erhel, B. Hamlat, and T. Migot, A projected dynamical system approach to mineral precipitation-dissolution reactions in geochemistry, Preprint.

[35] J. Erhel and T. Migot, Characterizations of solutions in geochemistry: existence, uniqueness, and precipitation diagram, Computational Geosciences, 23 (2019), pp. 523–535.

[36] J. Erhel and S. Sabit, Analysis of a global reactive transport model and results for the MoMas benchmark, Mathematics and Computers in Simulation, 137 (2017), pp. 286–298.

[37] J. Erhel, S. Sabit, and C. de Dieuleveult, Solving partial differential algebraic equations and reactive transport models, Computational Science, Engineering and Technology Series.

[38] A. F. Filippov, Differential equations whose right-hand is discontinuous on intersecting surfaces, (1979).

[39] ——, Differential Equations with Discontinuous Righthand Sides, 1988.

[40] J. C. Friedly and J. Rubin, Solute transport with multiple equilibrium-controlled or kinetically-controlled chemical reactions, Water resources research, 28 (1992), pp. 1935–1953.

[41] N. Guglielmi and E. Hairer, Classification of hidden dynamics in discontinuous dynamical systems, SIAM J. APPLIED DYNAMICAL SYSTEMS, (2015).

[42] F. Haeberlein, Time Space Domain Decomposition Methods for Reactive Transport, Application to CO2 Geological Storage, theses, Université Paris-Nord - Paris XIII, Oct. 2011.

[43] B. Hamlat, J. Erhel, A. Michel, and T. Faney, Kinetic models with pure phases, Preprint.

[44] J. Hoffmann, Reactive transport and mineral dissolution/precipitation in porous media : efficient solution algorithms, benchmark computations and existence of global solutions, PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2010.

[45] J. Hoffmann, S. Kräutle, and P. Knabner, A general reduction scheme for reactive transport in porous media, Computational Geosciences, 16 (2012), pp. 1081–1099.

[46] J. Hoffmann, S. Kräutle, and P. Knabner, Existence and uniqueness of a global solution for reactive transport with mineral precipitation-dissolution and aquatic reactions in porous media, SIAM Journal on Mathematical Analysis, 49 (2017), pp. 4812–4837.

[47] A. Jourani and E. Vilches, Moreau-Yosida regularization of state-dependent sweeping processes with nonregular sets, Journal of Optimization Theory and Applications, 173 (2017), pp. 91–116.

[48] C. Kanzow, N. Yamashita, and M. Fukushima, New ncp-functions and their properties, Journal of Optimization Theory and Applications, 94 (1997), pp. 115–135.

[49] P. Knabner, C. J. van Duijn, and S. Hengst, An analysis of crystal dissolution fronts in flows through porous media. Part 1: Compatible boundary conditions, Advances in Water Resources, 18 (1995), pp. 171–185.

[50] S. Kräutle, General multi-species reactive transport problems in porous media: Efficient numerical approaches and existence of global solutions, habilitation thesis, University of Erlangen-Nuremberg, Department of Mathematics, Erlangen, 2008.

[51] S. Kräutle, Existence of global solutions of multicomponent reactive transport problems with mass action kinetics in porous media, Journal of Applied Analysis and Computation, (Volume 1, Number 4, November 2011).

[52] S. Kräutle, J. Hodai, and P. Knabner, Robust simulation of mineral precipitation-dissolution problems with variable mineral surface area, submitted, (2019).

[53] K. Kumar, I. S. Pop, and F. A. Radu, Convergence analysis for a conformal discretization of a model for precipitation and dissolution in porous media, Numerische Mathematik, 127 (2014), pp. 715–749.

[54] A. C. Lasaga, Kinetic Theory in the Earth Sciences, 1997.

[55] A. Lauser, C. Hager, R. Helmig, and B. Wohlmuth, A new approach for phase transitions in miscible multi-phase flow in porous media, (2010).

[56] A. M. Leal, D. A. Kulik, W. R. Smith, and M. O. Saar, An overview of computational methods for chemical equilibrium and kinetic calculations for geochemical and reactive transport modeling, Pure and Applied Chemistry, 89 (2017), pp. 597–643.

[57] P. C. Lichtner, C. I. Steefel, and E. H. Oelkers, Reactive transport in porous media, Walter de Gruyter GmbH & Co KG, 2019.

[58] E. W. Lund, Guldberg and waage and the law of mass action, Journal of Chemical Education, (1965), p. 548–550.

[59] O. L. Mangasarian, Nonlinear programming, Society for Industrial and Applied Mathematics, 1994.

[60] M. M. Morel and J. G. Hering, Principles and Applications of Aquatic Chemistry, 1993.

[61] A. Nagurney and D. Zhang, Projected dynamical systems and variational inequalities with applications, Springer Science & Business Media, 1996.

[62] E. Nourtier-Mazauric, Modélisation géochimique et numérique des interactions entre des solutions solides et une solution aqueuse. Extension du logiciel de réaction-transport Archimède et application à la diagenèse des réservoirs, PhD thesis, 2003. Thèse de doctorat dirigée par Guy, Bernard et Garcia, Daniel Sciences de la Terre et de l'Environnement Saint-Etienne, EMSE 2003.

[63] L. Qi and D. Sun, A survey of some nonsmooth equations and smoothing Newton methods, (April 20, 1998; Revised August 13, 1998).

[64] R. T. Rockafellar and R. J.-B. Wets, Variational analysis, vol. 317, Springer Science & Business Media, 2009.

[65] J. Rubin, Transport of reacting solutes in porous media: Relation between mathematical nature of problem formulation and chemical nature of reactions, Water Resources Research, 19 (1983), pp. 1231–1252.

[66] S. Sabit, Numerical methods for reactive transport, PhD thesis, Université Rennes 1, 2014.

[67] L. F. Shampine and M. W. Reichelt, The matlab ode suite, SIAM J. Sci. Comput., 18 (1997), pp. 1–22.

[68] E. Tillier, Couplage réactions-transport pour la modélisation et la simulation du stockage géologique de CO2, PhD thesis, 2007. Thèse de doctorat dirigée par Eymard, Robert Mathématiques appliquées Université de Marne-la-Vallée 2007.

[69] V. I. Utkin, Sliding modes in control and optimization, 1992.

[70] C. J. van Duijn and P. Knabner, Travelling wave behaviour of crystal dissolution in porous media flow, European Journal of Applied Mathematics, 8 (1997), pp. 49–72.

[71] C. J. van Duijn, P. Knabner, and R. J. Schotting, An analysis of crystal dissolution fronts in flows through porous media part 2: incompatible boundary conditions, Advances in water resources, 22 (1998), pp. 1–16.

[72] C. J. van Duijn and I. S. Pop, Crystal dissolution and precipitation in porous media: pore scale analysis, Journal für die reine und angewandte Mathematik (Crelles Journal), 2004 (2004), pp. 171–211.

UNIVERSITÉ DE RENNES 1

Titre : Modélisation mathématique de réactions cinétiques multiphasiques en géochimie

Mot clés : cinétique chimique, EDO discontinue, inclusion différentielle, système dynamique projeté, théorie de Filippov, simulation numérique, analyse d'existence.

Résumé : Cette thèse concerne la modélisation mathématique des réactions cinétiques comprenant des phases pures. Dans le premier chapitre, un modèle de type EDOs discontinues pour la cinétique avec apparitions et disparitions d'espèces pour un nombre quelconque de minéraux est proposé. Une version régularisée du modèle permet de prouver la positivité et l'existence. Une analyse explicite plus approfondie dans le cas contenant une espèce réactive intermédiaire est menée. Dans le deuxième chapitre, une reformulation du modèle de cinétique chimique utilisant la théorie de Filippov est proposée. Une preuve de l'existence et de la positivité des solutions est réalisée. De plus, dans le cas des surfaces de discontinuité de codimension 1, une étude des configurations des champs fournit un résultat d'unicité et de caractérisation des trajectoires. Dans le troisième chapitre, un modèle de cinétique chimique de type systèmes dynamiques projetés est proposé. Une analyse de l'existence des solutions de ce modèle, des liens avec d'autres types de formulations et une méthode de résolution numérique adaptée sont proposés. Enfin, une illustration des résultats numériques obtenus est réalisée pour des systèmes de cinétique chimique.

Title: Mathematical modeling of multiphase kinetic reactions in geochemistry

Keywords: chemical kinetics, discontinuous ODEs, differential inclusions, projected dynamical system, Filippov's theory, numerical simulations, existence analysis.

Abstract: This thesis focuses on the modeling chemical kinetics for reactions involving pure phases. In the first chapter, a discontinuous ODEs model for reactions with appearance and disappearance of species for any number of minerals is proposed. A regularized version of the model can prove positivity and existence. An explicit analysis in the case containing an intermediate reactive species is investigated. In the second chapter, a reformulation of the chemical kinetics model using Filippov's theory is proposed. A proof of the existence and the positivity of the solutions is given. In addition, in the case of discontinuity surfaces of codimension 1, a study of the configurations of the vector fields provides a result of uniqueness and a characterization of the trajectories. In the third chapter, a model of chemical kinetics of the projected dynamical system is proposed. An analysis of the existence of solutions of this model, links with other types of formulations and an adapted numerical resolution method are provided. An illustration of the numerical results obtained is made for chemical kinetic systems.