



HAL
open science

Modèles couplés en milieux poreux : transport réactif et fractures

Laila Amir

► **To cite this version:**

Laila Amir. Modèles couplés en milieux poreux : transport réactif et fractures. Mathématiques [math]. Université Paris Dauphine - Paris IX, 2008. Français. NNT : . tel-00373688

HAL Id: tel-00373688

<https://theses.hal.science/tel-00373688>

Submitted on 7 Apr 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS DAUPHINE

EDDIMO

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

DOCTEUR EN MATHÉMATIQUES APPLIQUÉES

préparée au centre de recherche **INRIA (Rocquencourt)**
& **ITASCA Consultants S.A.S (Ecully)**

présentée et soutenue publiquement
par

Laila AMIR

le 18 Décembre 2008

Titre:

**Modèles couplés en milieux poreux :
transport réactif et fractures**

Directrice de thèse: **Jean E. ROBERTS**

Co-directeur de thèse: **Michel KERN**

Jury

M. Guy CHAVENT ,	Professeur à l'université Paris Dauphine	Président du jury
M. Brahim AMAZIANE ,	Maître de conférences à l'université de Pau	Rapporteur
M ^{me} . Jocelyne ERHEL ,	Directrice de recherche à l'INRIA de Rennes	Rapporteur
M. Daniel BILLAUX ,	Président de la société ITASCA Consultants	Examineur
M. Michel KERN ,	Chargé de recherche à l'INRIA Rocquencourt	Co-directeur de thèse
M ^{me} . Jean E. ROBERTS ,	Directrice de recherche à l'INRIA Rocquencourt	Directrice de thèse

Résumé : Cette thèse porte sur la simulation numérique de modèles couplés pour l'écoulement et le transport dans les milieux poreux. Nous présentons une nouvelle méthode de couplage entre les réactions chimiques et le transport en utilisant une méthode de Newton-Krylov, et nous étudions également un modèle d'écoulement en milieu fracturé qui traite l'intersection des fractures par une méthode de décomposition de domaine. Ce travail est divisé en trois parties : la première partie contient une analyse de différents schémas numériques pour la discrétisation des problèmes d'advection-diffusion, notamment par une technique de séparation d'opérateurs, ainsi que leur mise en oeuvre informatique, dans un code industriel. La deuxième partie, qui est la contribution majeure de cette thèse, est consacrée à la modélisation et à l'implémentation d'une méthode de couplage globale pour le transport réactif. Le système couplé transport-chimie est décrit, après discrétisation en temps, par un système d'équations non linéaires. La taille du système sous-jacent, à savoir le nombre de points de grille multiplié par le nombre d'espèces chimiques, interdit la résolution du système linéaire par une méthode directe. Pour remédier à cette difficulté, nous utilisons une méthode de Newton-Krylov qui évite de former et de factoriser la matrice Jacobienne. Dans la dernière partie, nous présentons un modèle d'écoulement dans un milieu fracturé tridimensionnel, basé sur une méthode de décomposition de domaine, et qui traite l'intersection des fractures. Nous démontrons l'existence et l'unicité de la solution, et nous validons le modèle par des tests numériques.

Mots clés : Milieu poreux, écoulement, advection-diffusion, réaction chimique, couplage, Newton-Krylov, intersection des fractures, décomposition de domaine.

Abstract : This thesis deals with numerical simulation of coupled models for flow and transport in porous media. We present a new method for coupling chemical reactions and transport by using a Newton-Krylov method, and we also present a model of flow in fractured media, based on a domain decomposition method that takes into account the case of intersecting fractures. This study is composed of three parts : the first part contains an analysis, and implementation, of various numerical methods for discretizing advection-diffusion problems, in particular by using operator splitting methods. The second part is concerned with a fully coupled method for modeling transport and chemistry problems. The coupled transport-chemistry model is described, after discretization in time, by a system of nonlinear equations. The size of the system, namely the number of grid points times the number a chemical species, precludes a direct solution of the linear system. To alleviate this difficulty, we solve the system by a Newton-Krylov method, so as to avoid forming and factoring the Jacobian matrix. In the last part, we present a model of flow in 3D for intersecting fractures, by using a domain decomposition method. The fractures are treated as interfaces between subdomains. We show existence and uniqueness of the solution, and we validate the model by numerical tests.

Keywords : Porous media, flow, advection-diffusion, chemical reaction, coupling, Newton-Krylov, intersecting fractures, domain decomposition.

Remerciements

Cette thèse a été effectuée dans le cadre d'une convention CIFRE (Convention Industrielle de Formation par la Recherche) qui a réuni la société ITASCA Consultants et le groupe ESTIME de l'INRIA Rocquencourt.

Ce travail de recherche n'aurait pu aboutir sans une réelle collaboration et un échange des idées entre tous ceux qui y ont participé ; je tiens ici les remercier.

En premier lieu, je tiens à exprimer ma vive reconnaissance à mes directeurs de thèse Madame Jean. E. Roberts, Directrice de recherche à l'INRIA Rocquencourt et Monsieur Michel Kern Chargé de recherche à l'INRIA Rocquencourt. Ils se sont montrés toujours disponible et m'ont constamment aidé à y voir plus clair. Ils ont fait preuve à la fois d'une grande patience, gentillesse, et d'un esprit responsable, critique et rigoureux. Ils m'ont vivement encouragé pour participer à de nombreuses conférences internationales et nationales qui m'ont enrichie et donnée de l'assurance. Je n'oublierai jamais leur sympathie et l'amitié que j'ai trouvé en eux, je ne sais pas comment je peux leur exprimer toute ma gratitude, merci.

Mes remerciements vont également à Monsieur Daniel Billaux, Président de la société Itasca Consultants S.A.S à Ecully, pour la confiance qu'il m'a accordé en acceptant le financement de ce travail. Je le remercie également d'avoir accepté de faire partie de mon jury de thèse. J'adresse aussi mes remerciements à Madame Caroline DARCEL, Ingénieur en hydrogéologie pour ses conseils scientifiques avisées, ainsi je tiens à remercier toute l'équipe ITASCA plus particulièrement Madame Christine ARNAUD pour leur accueil chaleureux dans l'entreprise.

Mes plus vifs remerciements vont à Monsieur Jérôme JAFFRE, Directeur de Recherche à l'INRIA Rocquencourt et chef du projet ESTIME, pour la confiance qu'il m'a accordée en m'associant aux recherches de son équipe. Son suivi précieux, ses qualités scientifiques et humains ont largement contribué à l'aboutissement de cette thèse. Tous mes remerciements vont aussi aux membres de l'équipe avec qui j'ai eu un très grand plaisir de travailler.

Mes sincères reconnaissances vont aux personnes qui ont accepté d'être rapporteurs de cette thèse, et ainsi évaluer mon travail : Madame Jocelyne ERHEL, Directrice de recherche à l'INRIA de Rennes et Monsieur Brahim AMAZIANE, Maître de conférences à l'Université de Pau et de Pays de l'Adour. Je remercie également Monsieur Guy CHAVENT, Professeur à l'université Paris IX Dauphine de m'avoir fait l'honneur de juger mon travail de thèse.

Merci à tous les gens que j'ai eu le plaisir de côtoyer durant ces quelques années, à tous mes amis pour leur soutien morale et leur aide, plus particulièrement à Ithrei, Jalila, Tayeb, Bruno, ...

Mes plus profonds remerciements vont à mes parents, à mes frères et soeurs. Tout au long de mon cursus, ils m'ont toujours soutenu, encouragé et aidé. Ils ont su me donner toutes les chances pour réussir. à qui je doit tout.

Je passe ensuite une dédicace spéciale à ma chère petite soeur Mouna, pour tout ce qu'elle a fait et pour le soutien qu'elle m'a apporté durant les deux dernières années.

Table des matières

Table des matières	vii
1 Introduction générale	1
1 Motivations	2
1.1 Stockage de déchets radioactifs en formation profonde	2
1.2 Problème de stockage de gaz carbonique	3
1.3 Hétérogénéité des milieux poreux	3
2 Sujets abordés dans la thèse	6
2.1 Chapitre 2 : problème d’advection diffusion	7
2.2 Chapitres 3 et 4 : transport des contaminants chimiques	7
2.3 Chapitre 5 : Ecoulement dans l’intersection des fractures	8
2 Résolution du problème de transport	9
1 Introduction	9
2 Équations de l’écoulement et du transport	9
2.1 La loi de Darcy	10
2.2 L’équation de conservation de la masse	10
2.3 L’équation d’advection-diffusion	10
3 Choix d’une méthode numérique	11
4 Etude générale de la méthode de décomposition d’opérateurs	14
5 Étude des schémas numériques pour l’équation d’advection-diffusion en 1D	16
5.1 Étude du cas homogène	16
5.2 Étude du cas hétérogène	25
5.3 Etude de convergence numérique	32
6 Mise en oeuvre du transport par décomposition d’opérateurs en 3D dans le logiciel 3FLO	33
6.1 L’écoulement transitoire par la méthode mixte hybride	34
6.2 Transport : l’advection par VF et la diffusion par EFMH	42
6.3 Mise en oeuvre dans 3FLO pour la partie diffusive du transport	45
7 Validation numérique du transport avec la solution analytique	48
7.1 Cas test académique	48
7.2 La solution analytique	49
7.3 Comparaison des résultats	50

8	Conclusion	53
3	Résolution du modèle d'équilibre chimique	55
1	Modélisation des réactions chimiques	55
1.1	Réactions homogènes ou hétérogènes	56
1.2	Réactions cinétiques ou à l'équilibre	56
2	Modélisation des réactions d'équilibre	57
2.1	Espèces composantes et secondaires	57
2.2	Mise en équation	59
2.3	Le tableau de Morel	60
2.4	Existence et unicité pour le système d'équilibre	61
3	Résolution numérique du système d'équilibre	62
3.1	Exemples	63
4	Modélisation des réactions de sorption	65
4	Le couplage transport chimie	67
1	Introduction	67
2	Équations du transport réactif	69
2.1	Modèle de transport, cf chapitre 2	69
2.2	Équations chimiques, cf chapitre 3	69
2.3	Couplage transport et chimie	73
3	Formulations et algorithmes pour le couplage	74
3.1	Les différents méthodes de couplage	76
4	Approche globale basée sur la méthode Newton–Krylov	81
5	Résultats numériques	86
5.1	Échange d'Ions : exemple EX11 de PhreeqC	86
5.2	Benchmark GDR MoMaS : 1D easy	88
5	Theoretical and numerical study for flow in a porous medium with intersecting fractures	95
1	Introduction	95
2	Flow in a porous medium : problem with fractures	96
3	The weak formulation of the problem	100
4	Existence and uniqueness of the solution	101
5	The discrete formulation of the problem	107
6	Domain decomposition for dealing with fractures	109
7	Numerical results	111
7.1	Description the test cases	111
7.2	Performance of the preconditioner	115
8	Conclusion	117
	Conclusion générale	119

Perspectives	121
Bibliographie	125

Chapitre 1

Introduction générale

La migration d'un contaminant en solution est généralement le résultat de l'interaction de nombreux processus physiques, chimiques et biologiques. Les quatre principaux processus contrôlant le mouvement des contaminants en subsurface sont l'advection, la dispersion, le transfert de masse entre différentes phases et les réactions (les phénomènes de sorption, les partitions liquide-liquide ou la volatilisation). Les réactions correspondent à tous les processus modifiant la nature physico-chimique du contaminant.

La modélisation mathématique de ces phénomènes conduit à un problème couplé. Pour ce qui est des méthodes numériques, les problèmes de couplage sont le plus souvent à l'origine de beaucoup de difficultés. On peut en effet distinguer deux grandes familles de problèmes de couplage :

- le couplage de différents processus qui évoluent dans la même région de l'espace. Ainsi les couplages chimie-transport, écoulement-transport ou le couplage thermo-hydro-mécanique. Prenons l'exemple du couplage écoulement-transport, le champ de vitesse est transmis du code écoulement vers le code transport et inversement, la concentration est transmise du code transport vers le code écoulement.
- le couplage de processus qui évoluent dans des régions différentes de l'espace, caractérisés par leur très forte hétérogénéité venant de la large variabilité des propriétés hydrauliques du milieu. On peut donner comme exemples le couplage entre un aquifère où le transport est dominé par l'advection et une région où le processus de diffusion est dominant, ou bien le couplage entre un centre de stockage et le milieu géologique, ou encore le couplage entre des roches et les fractures qui les traversent.

Ces deux types de couplage présentent des difficultés numériques que les logiciels actuels prennent plus ou moins bien en compte. En général, la solution utilisée consiste à faire des simulations avec un seul modèle incluant tous les aspects phy-

siques, ce qui rend les calculs coûteux et mène rapidement aux limites des possibilités actuelles des calculateurs, notamment pour les simulations en trois dimensions.

Nous nous intéressons principalement dans cette thèse au couplage transport chimie sous l'hypothèse d'équilibre chimique. Et nous nous sommes particulièrement préoccupés de problèmes liés à l'hétérogénéité des milieux, en étudiant également l'écoulement dans les milieux poreux tridimensionnels fracturés, en prenant en compte l'intersection des fractures. Il s'agit d'un couplage géométrique fractures 2D-milieu 3D, intersection 1D-fractures 2D.

1 Motivations

1.1 Stockage de déchets radioactifs en formation profonde

La plupart des activités humaines produisent des déchets dont certains restent potentiellement dangereux pendant de très longues durées (des polluants chimiques, des déchets radioactifs). Les déchets radioactifs sont considérés (à vie courte) quand au bout de trente ans, plus de la moitié de la radioactivité a disparu. Dans le cas contraire les déchets sont considérés à vie longue (de centaines d'années à plusieurs milliers d'années).

Mais dans tous les cas, la gestion de ces déchets dangereux appelle bien évidemment au devoir de nous assurer que les déchets que nous produisons aujourd'hui n'imposeront à nos descendants aucune nuisance inacceptable, ni demain, ni à plus long terme.

La gestion des déchets radioactifs consiste essentiellement à les emprisonner sous forme solide et stable à l'intérieur de structures étanches, elles-mêmes durables, de façon à non seulement se protéger de leurs rayonnements, mais aussi à éviter leur retour et leur dispersion dans la biosphère.

C'est au regard de cette exigence éthique que plusieurs organismes analysent le problème des déchets radioactifs. En particulier, L'ANDRA (l'Agence Nationale pour la gestion des Déchets RAdioactifs), est un organisme public chargé de l'étude et de la gestion présente et future des déchets nucléaires. Cela concerne aussi le CEA (Commissariat à l'Énergie Atomique) et la COGEMA (Compagnie générale des matières nucléaires).

D'autres organismes publics de contrôle sont impliqués : la DSNI (Direction pour la Sûreté des Installations Nucléaires), l'ASN (Agence pour la Sûreté Nucléaire) ou l'IRSN (Institut de Radioprotection pour la Sûreté Nucléaire). Ils interviennent dans de nombreuses étapes de l'industrie nucléaire : dans les centrales, les réacteurs nucléaires, dans le transport et la surveillance des déchets, dans le contrôle du respect des normes et des dispositifs de sécurité.

Dans le cadre du stockage profond, la réalisation de simulations numériques

fiables est indispensable, puisque c'est le seul moyen de prédire le comportement des déchets en sous-sol sur des durées de plusieurs milliers d'années. Il est notamment nécessaire d'effectuer des simulations du transport de radionucléides en milieu souterrain afin de vérifier dans quelles conditions la proportion de radio éléments qui ressortent dans la biosphère reste en dessous des normes légales.

1.2 Problème de stockage de gaz carbonique

Le changement climatique apparaît comme l'un des principaux dangers qui menacent notre environnement. Le CO_2 est cité comme l'un des facteurs de risque majeurs, puisque son augmentation serait responsable de la tendance au réchauffement climatique, et pourrait avoir des conséquences dramatiques si aucune mesure n'est prise. Le transport, la capture et le stockage de CO_2 constituent un défi important pour lutter contre l'accumulation de CO_2 dans l'atmosphère. Une autre solution est de réduire la consommation d'énergie, ou de mettre en oeuvre de combustibles ou de carburants émettant moins de CO_2 par unité d'énergie produite.

Le stockage du CO_2 est confronté à des enjeux techniques et réglementaires, il doit notamment démontrer qu'il ne cause aucun dommage à l'environnement local. Trois solutions de stockage sont envisagées pour assurer le confinement du CO_2 sur le long terme :

- les anciens réservoirs d'hydrocarbures liquides ou gazeux,
- les aquifères salins profonds,
- les veines de charbon non exploitées.

L'IFP (l'Institut Français de Pétrole) travaille sur ces 3 axes, en mettant au point des logiciels industriels permettant de modéliser et de gérer les stockages. Il s'agit notamment des logiciels qui modélisent les voies possibles de migration du CO_2 dans le sous-sol, en fonction de la structure géologique du stockage et des interactions géochimiques entre le CO_2 et les structures minérales rencontrées.

Différents phénomènes sont mis en enjeu lors de l'injection de CO_2 :

- Le transport : advection, diffusion et dispersion
- Les réactions chimiques : dissolution du CO_2 dans l'eau, réactions entre l'eau et la roche.

Ces phénomènes se produisent à des échelles de temps très différents (réactions chimiques instantanées, réactions cinétiques lentes, transport à faible vitesse). La modélisation de l'ensemble de ces phénomènes revient à coupler des modèles d'écoulement multiphasiques et des modèles géochimiques.

1.3 Hétérogénéité des milieux poreux

Les milieux poreux sont généralement hétérogènes. Parmi les raisons citons la présence de fractures ou les réactions chimiques hétérogènes (différents états de la

matière). La prise en compte de l'hétérogénéité du milieu dans les modèles d'écoulement et de transport est un problème majeur de l'hydrogéologie.

a) Présence des fractures

L'hétérogénéité est souvent associée à la présence de fractures qui, malgré le faible volume qu'elles occupent dans l'aquifère, contrôlent l'essentiel des chemins de circulation. Les fractures forment les chemins de circulation préférentiels des fluides souterrains. Elles augmentent la vitesse d'écoulement des fluides ce qui favorise les transferts de masse et d'énergie.



FIG. 1.1: une roche fracturé

On peut distinguer deux types de fractures :

- les grandes fractures qui représentent des discontinuités géologiques. Elles sont en général beaucoup plus perméables que le milieu ambiant, devenant des canaux privilégiés pour l'écoulement. Mais elle peuvent être au contraire peu perméables, représentant des barrières géologiques.
- Le deuxième type de fracture est celui de petites fractures qui apparaissent en grand nombre dans le milieu, formant un réseau de fissures, dont les localisations précises sont difficiles à déterminer.

Vu leur rôle hydraulique, il est important de prendre en compte les fractures dans un modèle numérique d'écoulement ou de transport dans les milieux poreux.

La modélisation de l'écoulement ou du transport dans les milieux contenant les réseaux de fissures est souvent traitée par un modèle de double porosité [29], [30]. Pour la modélisation dans les milieux contenant les grandes fractures, une méthode classique s'appuie sur un raffinement local du maillage, et une autre approche est

d'utiliser une méthode inspirée de la décomposition de domaines. La fracture, supposée sans épaisseur mais avec des propriétés physiques propres, est remplacée par une condition de transmission. En éliminant la pression dans les sous-domaines ainsi définis, on se ramène à un problème d'interface de type non-standard.

b) Réactions hétérogènes

Une autre raison de l'hétérogénéité, les réactions chimiques dans le milieu poreux. De ce point de vue, les réactions les plus favorisées sont les réactions homogènes qui se déroulent au sein d'une seule phase (liquide : solution, mélange de liquides miscibles ou gazeuse). En effet, dans ces cas, les molécules des réactifs sont dans une même phase et peuvent donc facilement entrer en contact pour réagir. Dans le cas de systèmes hétérogènes, c'est-à-dire de réactions entre un solide et un gaz, un solide et un liquide, un liquide et un gaz, ou entre deux liquides non miscibles, la réaction ne peut avoir lieu qu'aux surfaces de séparation des phases (interfaces). Les réactions qui se produisent à l'interface entre le fluide et la roche sont principalement les réactions de sorption.

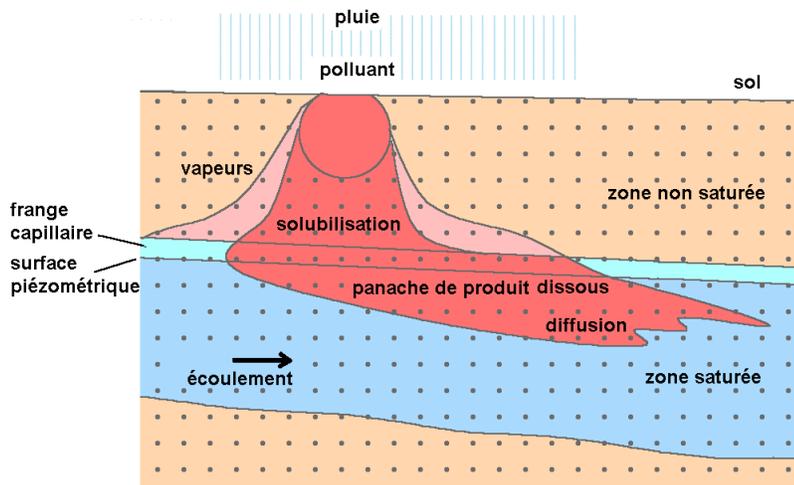


FIG. 1.2: pollution d'un milieu souterrain

Dans le milieu souterrain, l'évolution des polluants est très lente, comparative-ment à ce qui se passe en surface. Il s'écoule des mois, des années, ou même des dizaines d'années, entre le début de la pollution et sa mise en évidence.

Le temps de transfert d'un polluant chimique à la nappe varie de quelques jours à plusieurs années. Il dépend principalement de l'épaisseur de la zone non saturée, de la perméabilité du réservoir et des caractéristiques du polluant. Les terrains à perméabilité d'interstices (sables, grès) connaissent un temps de transfert très court.

Le soluté traverse la zone non saturée, et d'abord le sol, caractérisée par la présence d'oxygène, de minéraux argileux et matières organiques... Les réactions d'adsorption/désorption entraînent un retard dans le transfert du polluant. Ces deux phénomènes sont favorisés par la présence d'argiles, oxydes et hydroxydes, et par les constituants organiques. L'adsorption prédomine sur la désorption. Les métaux sous forme ionique sont retenus par échanges d'ions ou adsorption, par les argiles, acides humiques et hydroxydes.

En zone saturée, la vitesse de propagation du soluté n'est pas identique à celle de l'eau souterraine. Elle dépend du type de polluant, en particulier de sa viscosité, et de sa concentration (phénomène de diffusion). Le débit de la nappe et les fluctuations de la surface piézométrique jouent également un rôle.

Transfert dans le sol : Grâce à ses propriétés d'adsorption et d'échange, dues à la présence de colloïdes minéraux et organiques, le sol peut retenir un grand nombre de substances très diverses. Les polluants peuvent être piégés dans les pores du sol.

Transfert dans la zone non saturée : Le fluide polluant migre d'abord verticalement dans le milieu non saturé entre la surface du sol et la nappe, laissant dans son sillage des terrains imprégnés à une concentration proche de la saturation. Suivant sa volatilité, le polluant a tendance à plus ou moins diffuser dans la phase gazeuse du milieu non saturé. C'est notamment le cas des solvants chlorés et des hydrocarbures aromatiques.

Evolution en milieu saturé : Dans l'aquifère, les mécanismes de transport du soluté sont complexes. Ils sont la conséquence de l'hétérogénéité du réservoir. La dispersion du soluté se fait à la verticale de la zone non saturée, puis selon un étalement latéral dans le sens de l'écoulement de l'eau souterraine (zone saturée). La vitesse d'écoulement est assez lente en aquifère homogène, mais peut être extrêmement rapide en milieu plus perméable.

On distingue les éléments chimiques solubles des insolubles, ils peuvent rester à la surface de la nappe où ils s'évalent. La concentration des polluants minéraux dépend de leur solubilité. Les substances organiques solubles sont maintenues en solution sous forme moléculaire par des liaisons hydrogènes. Les produits organiques peuvent être totalement détruits dans l'eau, ou transformés en produits non toxiques, par des réactions chimiques.

2 Sujets abordés dans la thèse

Nous avons mené des études principalement dans trois directions qui ont toutes en commun la simulation des écoulements et du transport en milieu poreux. Cette thèse est donc composée de trois parties. La première partie contient une mise en

oeuvre de différents schémas numériques pour la résolution du problème d'advection-diffusion. La deuxième partie est consacrée à la modélisation et à l'implémentation d'une méthode de couplage globale entre le transport et la chimie, basée sur une méthode de Newton-Krylov. Dans la dernière partie, on présente un modèle d'écoulement pour les fractures qui s'intersectent dans les milieux poreux. Les fractures sont traitées comme des interfaces entre des sous domaines. Plus précisément, nous résumons le contenu des différentes parties :

2.1 Chapitre 2 : problème d'advection diffusion

Cette partie porte sur la résolution de l'équation de transport en milieu poreux saturé. L'hydrodynamique des écoulements est décrite par un système d'équations composé de la loi de Darcy et de l'équation de conservation de la masse. Le transport des polluants est régi par l'équation classique de diffusion-advection. A cette équation de transport, on peut adjoindre différents termes décrivant les mécanismes chimiques.

La résolution de l'équation du transport nécessite de connaître la vitesse de l'écoulement. Cette vitesse est approchée en résolvant les deux équations d'écoulement par la méthode des éléments finis mixtes hybride, indépendamment du transport. Ensuite, cette vitesse est utilisée pour simuler le transport des polluants.

Pour la résolution de l'équation du transport, la technique de séparation d'opérateurs est utilisée pour discrétiser l'équation de diffusion-advection. Le terme diffusif est approché en utilisant une méthode d'éléments finis mixtes hybrides, et le terme convectif est approché par une méthode de volumes finis.

Dans un premier paragraphe nous rappelons les équations d'écoulement et du transport, ensuite nous étudions trois schémas pour la discrétisation de l'équation du transport en 1D dans les cas homogène et hétérogène. La suite du chapitre se concentre sur la résolution des équations d'écoulement en régime transitoire ou permanent et la résolution de l'équation de transport en 3D par la méthode de séparation d'opérateurs. Nous détaillons la mise en oeuvre de l'écoulement transitoire et du transport dans un code industriel. Toutes les méthodes numériques ont été validées par des solutions analytiques sauf dans le cas hétérogène où la solution numérique a été validée par une solution référence.

2.2 Chapitres 3 et 4 : transport des contaminants chimiques

Prédire la migration d'un radionucléide dans un milieu nécessite d'être capable de modéliser l'influence du transport sur la chimie et réciproquement. Dans certains cas, ce couplage peut se modéliser simplement par un terme de retard relié au coefficient de distribution introduit dans l'équation du transport et qui prend en compte la rétention aux interfaces. Cependant, il est souvent nécessaire de modéliser explicitement et simultanément la réactivité chimique et le transport par l'utilisation d'outils numériques dédiés : les codes couplés chimie-transport.

Il existe différentes types d'approche pour résoudre le transport réactif. La plupart des modèles effectuent le couplage chimie-transport selon deux grandes catégories, les approches globales (GIA, Global Implicit Approach) et les approches séquentielles, SNIA (Sequentiel Non Iterative Approach) et SIA (Sequentiel Iterative Approach).

Le transport de solutés réactifs en milieu poreux est un problème difficile du point de vue numérique. Il faut d'une part coupler le modèle de transport et le modèle de chimie. D'autre part, il faut résoudre de manière robuste le problème de l'équilibre chimique, qui est un système d'équations algébriques. Nous étudions d'abord la résolution des équations d'équilibre, sans prendre en compte les réactions de précipitation-dissolution. Les réactions de précipitation-dissolution présentent une difficulté particulière par leur caractère à seuil, donc non-différentiable. Nous examinons les méthodes classiques de couplage, en les réinterprétant à la lumière de l'analyse numérique, ce qui permet de suggérer des améliorations. Pour résoudre le problème couplé, on propose d'utiliser une (variante de la) méthode de Newton, en gardant à l'esprit qu'il n'est pas possible de former la matrice jacobienne du système couplé. Une solution possible est de résoudre le système linéaire à chaque itération par un solveur linéaire itératif, par exemple GMRES. On parle d'une méthode de Newton-Krylov. Nous montrons comment formuler le problème couplé transport-chimie dans ce formalisme, de manière à conserver la séparation entre les modules de transport et de chimie. Nous illustrons la méthode sur des exemples, dont un benchmark proposé par le groupe MoMaS.

2.3 Chapitre 5 : Ecoulement dans l'intersection des fractures

Nous nous intéressons à la modélisation de l'écoulement d'un fluide monophasique dans un milieu poreux fracturé, avec des fractures qui s'intersectent (c'est le cas des grandes fractures qui nous intéresse ici), en utilisant les méthodes de décomposition de domaine. Les fractures sont traitées comme des interfaces entre les sous-domaines. Nous démontrons l'existence et l'unicité de la solution, et pour la validation numérique, nous présentons les résultats obtenues sur un cas test académique en 3D.

Chapitre 2

Résolution du problème de transport

1 Introduction

L'objet de ce chapitre est la simulation de l'écoulement et du transport en milieu poreux saturé. L'hydrodynamique des écoulements est décrite par un système d'équations composé de la loi de Darcy et de l'équation de conservation de la masse. Le transport des polluants est régi par l'équation classique de diffusion-advection. A cette équation de transport, on peut adjoindre différents termes décrivant les mécanismes chimiques.

La résolution de l'équation du transport nécessite de connaître la vitesse de l'écoulement. Cette vitesse est approchée en résolvant les deux équations d'écoulement par la méthode des éléments finis mixtes hybride, indépendamment du transport. Ensuite, cette vitesse est utilisée pour simuler le transport des polluants.

Pour la résolution de l'équation du transport, la technique de séparation d'opérateurs est utilisée pour discrétiser l'équation de diffusion-advection. Le terme dispersif est approché en utilisant une méthode d'éléments finis mixtes hybrides, et le terme advectif est approché par une méthode de volumes finis.

Dans un premier paragraphe nous rappelons les équations d'écoulement et du transport, ensuite nous étudions trois schémas pour la discrétisation de l'équation du transport en 1D dans les cas homogène et hétérogène. La suite du chapitre se concentre sur la résolution des équations d'écoulement en régime transitoire ou permanent et la résolution du transport en 3D par la méthode de séparation d'opérateurs. Les méthodes numériques sont validées par des solutions analytiques, et mises en oeuvre dans un code industriel.

2 Équations de l'écoulement et du transport

L'écoulement d'un fluide dans un milieu poreux est régi par deux lois principales : la loi de Darcy et l'équation de conservation de la masse.

2.1 La loi de Darcy

Les causes de déplacement d'un fluide en milieux poreux sont les gradients de pression et la gravité. La loi de Darcy exprime la vitesse de filtration en fonction de la charge hydraulique. Pour un fluide incompressible, la loi de Darcy s'écrit :

$$u = -\frac{k\rho g}{\mu} \nabla \left(\frac{p}{\rho g} + z \right) = -\mathcal{K} \nabla h, \quad (2.1)$$

- u le vecteur vitesse de Darcy, $[LT^{-1}]$,
- k la perméabilité intrinsèque du milieu, $[L^2]$,
- ρ la masse volumique de l'eau, $[ML^{-3}]$,
- g la gravité, $[LT^{-2}]$,
- z la cote verticale, $[L]$,
- p la pression
- $h = \frac{p}{\rho g} + z$ la charge hydraulique, $[L]$,
- \mathcal{K} le coefficient ou le tenseur de perméabilité, $[LT^{-1}]$.

En trois dimensions, le tenseur de perméabilité \mathcal{K} est une matrice symétrique éventuellement pleine. Ce tenseur peut se réduire à ses composantes diagonales lorsqu'on se place dans le repère de coordonnées dont les axes sont les directions pour lesquelles l'écoulement est effectivement parallèle au gradient de la charge (de Marsily, [28]).

2.2 L'équation de conservation de la masse

L'équation de conservation de la masse exprime le principe de la conservation de masse d'un fluide en mouvement, dans un volume élémentaire (on parle de Volume Élémentaire Représentatif pour pouvoir faire un modèle continu d'un milieu hétérogène à l'échelle microscopique), la variation de la masse de fluide durant une unité de temps est égale à la somme algébrique des flux massiques traversant la surface du volume considéré. L'équation de bilan de masse s'écrit sous la forme générale :

$$S \frac{\partial h}{\partial t} + \operatorname{div} u = \tilde{f}, \quad (2.2)$$

où :

- S le coefficient d'emménagement $[L^{-1}]$,
- \tilde{f} le terme source par unité de volume $[T^{-1}]$.

Si la charge ne varie pas avec le temps, c'est le cas du régime permanent, l'équation (2.2) se réduit à la forme :

$$\operatorname{div} u = \tilde{f}. \quad (2.3)$$

2.3 L'équation d'advection-diffusion

Le transport des polluants est la superposition de plusieurs phénomènes hydrodynamiques (advection, dispersion, ..), chimiques (adsorption, précipitation,..) qui

peuvent influencer sur le déplacement. Le transport d'un polluant non réactif dû à l'advection et la dispersion est donné par :

$$\phi \frac{\partial c}{\partial t} + \operatorname{div}(-D\nabla c) + \operatorname{div}(uc) = f, \quad (2.4)$$

- c représente la concentration en solution [ML^{-3}],
- u la vitesse de filtration de l'écoulement, c.à.d la vitesse de Darcy [LT^{-1}], cf. section 2.1.
- ϕ la porosité du milieu [.] ,
- f le terme source [MT^{-1}],
- D est le tenseur de dispersion-diffusion donné en fonction de la vitesse u par :

$$D(u) = d_m I + |u| [d_l E(u) + d_t (I - E(u))], \quad \text{avec} \quad E_{ij}(u) = \frac{u_i u_j}{|u|^2}, \quad (2.5)$$

- d_m la diffusion moléculaire [$L^2 T^{-1}$],
- d_l la diffusion longitudinale [L],
- d_t la diffusion transversale [L],

La partie advective (2.6) de cette équation (2.4) représente le débit de l'advection, et c'est le déplacement moyen à la vitesse de Darcy,

$$\varphi_a = uc. \quad (2.6)$$

La partie diffusive (2.7) représente le débit de la diffusion, et c'est le déplacement dû à l'agitation moléculaire ou à l'hétérogénéité des vitesses microscopiques,

$$\varphi_d = -D\nabla c. \quad (2.7)$$

Le débit total φ est la somme de deux debits advectif et diffusif : $\varphi = \varphi_a + \varphi_d$.

On résout d'abord l'équation de Darcy et de conservation indépendamment du transport pour obtenir la vitesse. Ensuite, cette vitesse est utilisée pour simuler le transfert de polluant.

3 Choix d'une méthode numérique

On peut se voir obligé d'utiliser une solution numérique des équations d'écoulement et du transport plutôt que la solution analytique pour plusieurs raisons :

- Le domaine d'écoulement est borné par des limites qui influencent au cours du temps sur la solution recherchée. Généralement les solutions analytiques s'appliquent aux milieux infinis ou semi infinis.

- Le milieu est hétérogène : ses propriétés varient dans l'espace, tandis que les solutions analytiques supposent que milieu est homogène ou que la géométrie des hétérogénéités est très simple.
- Le problème est non linéaire et il n'existe pas de solution analytique.

Dans le cas où les solutions numériques sont nécessaires, il faut choisir une méthode bien adaptée pour la résolution d'un problème donné.

Différentes méthodes numériques, comme la méthode de Différences Finis (DF), de Volumes Finis (VF), des Éléments Finis de Galerkin (EFG), des Éléments Finis Mixtes (EFM), des Éléments Finis Discontinus (EFD), ou des Éléments Finis Mixte-Hybride (EFMH) peuvent être utilisées pour la résolution des équations d'écoulement en milieux poreux.

Les méthodes DF et VF permettent le calcul d'une charge (pression) ou d'une concentration (moyenne par élément pour les VF et aux noeuds pour les DF). Ces méthodes présentent des contraintes sur la discrétisation en espace. Les DF sont limitées à des mailles régulières simples (des carrés réguliers, des rectangles). Elles ne sont pas adaptées à des domaines à géométrie complexe. La méthode VF est plus flexible que celle des DF, cependant cette approche nécessite le respect des conditions de Delaunay sur la discrétisation spatiale du maillage triangulaires ce qui devient difficile pour un maillage en tétraèdres pour le cas tridimensionnel. De plus, ces deux méthodes ne permettent pas de traiter de manière commode le cas d'une perméabilité anisotrope d'orientation quelconque (tenseur non diagonal, plein).

Par ailleurs, les méthodes EFG standards permettent le calcul d'une charge ou d'une concentration aux noeuds des maillages. En pratique, on prend des triangles et des rectangles en deux dimension, et des tétraèdres ou des parallélépipèdes rectangles en trois dimension. Ceci permet de décrire d'une manière satisfaisante la géométrie du milieu ainsi que les hétérogénéités, et permet aussi de traiter toutes les directions d'anisotropie (tenseur de perméabilité plein). Cependant la méthode standard ne garantit pas la continuité des flux à l'interface entre les éléments, ce qui a pour effet de donner un champs de vitesse qui n'est pas physique.

La méthode des éléments finis mixtes (EFM) (présentée par la première fois par Meissner [62] et développée par d'autres [66], [31], [67], [15], [14], [16], [87], [70],...) apporte une réponse à ce type de problème en approchant simultanément soit la pression et la vitesse de Darcy, soit la concentration et le flux diffusif. Cette méthode aboutit à un système d'équations dont les inconnues sont les concentrations moyennes par maille et les flux diffusifs à travers les interfaces des éléments. La matrice associée à ce système est symétrique non définie positive et donc difficile à inverser.

La formulation EFMH (G. Chavent et J-E. Roberts [21], G. Chavent et J. Jaffré [20], J-E. Roberts et J.M. Thomas [71], F. Brezzi et M. Fortin [16]) permet de supprimer l'inconvénient en choisissant comme une nouvelle inconnue qui est un multiplicateur de Lagrange utilisé pour assurer la continuité du flux (de Darcy, ou diffusif). Ce multiplicateur est une approximation de la trace de la pression moyenne sur les arêtes ou les faces du maillage, on l'appelle trace de la pression ou de la concentration. La matrice associée au système mixte hybride est symétrique définie positive et ce type de système se résout de manière efficace avec des algorithmes tels que Cholesky, Gradient conjugué, ...

L'équation du transport peut être résolue par diverses méthodes : différences finies, élément finis, volumes finis, méthodes des particules, ... Le succès de ces méthodes repose souvent sur le respect du critère de Courant et de Péclet. Le nombre de Courant CFL et de Péclet de maille sont définis en une dimension par :

$$\text{CFL} = \frac{u\Delta t}{\phi\Delta x}, \quad \text{et} \quad \text{Pe} = \frac{u\Delta x}{D},$$

où Δx est le pas d'espace, Δt est le pas du temps, et en trois dimension par :

$$\text{CFL} = \frac{1}{2} \frac{\int_{\partial K} |u \cdot n| dx}{\phi_K |K|}, \quad \text{Pe} = \frac{1}{2} \frac{\int_{\partial K} |u \cdot n| dx}{\|D\|},$$

où K est un élément de discrétisation, et les limites à respecter sont : $\text{CFL} \leq 1$ et $\text{Pe} \leq 2$. Le respect du critère CFL assure une stabilité du schéma numérique, et le respect de la limite Pe permet d'éviter les oscillations en vérifiant le principe du maximum.

Les critères de Courant et de Péclet imposent des contraintes très strictes sur le pas de temps et la tailles des mailles, ce qui exige un temps de calcul et un besoin en place mémoire importants. Diverses méthodes ont été développées pour s'affranchir de ces critères. Une approche développée depuis plusieurs années (Siegel et al [80]) s'appuie sur une résolution séparée des opérateurs d'advection et de dispersion. La méthode de séparation d'opérateurs consiste à séparer l'équation du transport en différentes sous équations, qui sont traitées chacune de manière spécifique. Cette technique permet ainsi de traiter des processus différents comme l'advection, la diffusion, la chimie, avec des méthodes appropriés pour chaque type de phénomène.

La construction des méthodes de séparation d'opérateurs n'est pas unique, elle varie selon le choix des méthodes numériques utilisées pour les sous-problèmes obtenus. Douglas & Russell [32], et Bercovier et al [7], ont approché l'équation du transport en combinant la méthode des caractéristiques pour approcher la partie

advective et la méthode EF pour approcher la partie diffusive. D'autres combinaisons sont utilisées comme celle de la méthode EFMH et EFD, celle de la méthode EFMH et VF. Ces méthodes sont exposées dans [78], [47], [51], [23].

Pour le développement dans le code 3FLO (dont nous reparlerons plus tard), on a choisi ¹ la dernière combinaison pour résoudre le problème du transport. Pour la partie advective du transport (2.6), on a utilisé la méthode de volumes finis décentré en amont dans le but de préserver le bilan de masse localement, de donner une bonne représentation des fronts et de limiter la dispersion numérique. Pour la partie diffusive (2.7), on a utilisé la méthode des éléments finis mixtes hybrides qui conserve la masse localement en assurant la continuité du flux à travers les faces, et qui permet de traiter des tenseurs de dispersion pleins.

La condition de stabilité pour l'advection impose une contrainte particulièrement sévère sur le pas de temps quand l'advection est significative. Ceci se traduit par un coût de calcul très important, car le traitement implicite de la diffusion impose la résolution d'un système linéaire à chaque pas de temps. D'où vient l'idée de faire de petits pas de temps d'advection, contrôlés par une condition CFL, et des plus grands pas de temps de diffusion en utilisant la technique de décomposition d'opérateurs (splitting). Cette technique nous permet d'utiliser deux échelles différents du pas de temps. Pour cela, on divise le pas de temps de diffusion en M pas de temps d'advection contrôlé par la condition CFL. En procédant ainsi on pourra faire plusieurs pas de temps d'advection pour un pas de temps de diffusion, ce qui conduira à d'importants gains de temps de calcul, les pas de temps de diffusion-dispersion étant beaucoup plus coûteux que ceux d'advection puisqu'ils impliquent la résolution d'un système linéaire de grande taille.

4 Etude générale de la méthode de décomposition d'opérateurs

Cette partie concerne une étude générale de l'ordre et de l'erreur d'approximation numérique due à la décomposition d'opérateurs dans la méthode splitting.

¹ Mon travail porte principalement sur le couplage entre le transport et la chimie, j'ai eu besoin de mettre en oeuvre le code du transport en 3D dans le logiciel industriel 3FLO, et pour pouvoir choisir la bonne méthode à implémenter, j'ai fait d'abord des études du transport en 1D avec différentes schémas numériques, et une comparaison avec une solution analytique et une solution de référence a permis de favoriser le choix de la technique de décomposition d'opérateurs, en plus, c'est une méthode qui a marché dans divers applications, comme exemple (cf. P. Ackerer et al, benchmark Couplex [47]).

Soit l'équation suivante :

$$\begin{cases} \frac{\partial c}{\partial t} = (\mathcal{A} + \mathcal{B})c \\ c(x, 0) = c_0, \end{cases} \quad (2.8)$$

où \mathcal{A} et \mathcal{B} sont deux opérateurs différentiels.

La méthode de décomposition d'opérateurs est définie par deux étapes successives :

– Etape 1 de résolution du processus \mathcal{A} :

$$\begin{cases} \frac{\partial c^*}{\partial t} = \mathcal{A}c^* \\ c^*(x, 0) = c_0, \end{cases}$$

– Etape 2 de résolution de résolution du processus \mathcal{B} :

$$\begin{cases} \frac{\partial c^{**}}{\partial t} = \mathcal{B}c^{**} \\ c^{**}(x, 0) = c^*, \end{cases}$$

La valeur de c est approchée par c^{**} . L'analyse de l'erreur induite par la séparation des opérateurs est effectuée à partir des solutions exponentielles en utilisant un développement en série de Taylor.

La solution exacte de l'équation (2.8) s'écrit :

$$c(x, k) = e^{k(\mathcal{A}+\mathcal{B})}c(x, 0).$$

alors que la solution calculée est :

$$\begin{aligned} c^*(x, k) &= e^{k\mathcal{A}}c(x, 0), \\ c^{**}(x, k) &= e^{k\mathcal{B}}e^{k\mathcal{A}}c(x, 0), \end{aligned}$$

L'erreur de séparation d'opérateurs est donc :

$$c(x, k) - c^{**}(x, k) = (e^{k(\mathcal{A}+\mathcal{B})} - e^{k\mathcal{B}}e^{k\mathcal{A}}) c(x, 0), \quad (2.9)$$

et par utilisation de la série de Taylor,

$$\begin{aligned} c(x, k) &= c(x, 0) + k(\mathcal{A} + \mathcal{B})c(x, 0) + \frac{1}{2}k^2(\mathcal{A} + \mathcal{B})^2c(x, 0) + \dots \\ &= \left(I + k(\mathcal{A} + \mathcal{B}) + \frac{1}{2}k^2(\mathcal{A} + \mathcal{B})^2 \right) c(x, 0), \end{aligned}$$

et

$$\begin{aligned} c^{**}(x, k) &= \left(I + k\mathcal{B} + \frac{1}{2}k^2\mathcal{B}^2 + \dots \right) \left(I + k\mathcal{A} + \frac{1}{2}k^2\mathcal{A}^2 + \dots \right) c(x, 0), \\ &= \left(I + k(\mathcal{A} + \mathcal{B}) + \frac{1}{2}k^2(\mathcal{A}^2 + 2\mathcal{A}\mathcal{B} + \mathcal{B}^2) + \dots \right) c(x, 0) \end{aligned} \quad (2.10)$$

donc

$$c(x, k) - c^{**}(x, k) = \frac{1}{2}k^2 (\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}) c(x, 0) + O(k^3). \quad (2.11)$$

Ceci montre que l'erreur due à décomposition d'opérateurs est d'ordre 1 en temps, elle est nulle dans le cas où les deux opérateurs \mathcal{A} et \mathcal{B} commutent.

5 Étude des schémas numériques pour l'équation d'advection-diffusion en 1D

5.1 Étude du cas homogène

Le modèle

On s'intéresse dans cette partie à l'étude de l'équation d'advection-diffusion dans le cas monodimensionnel, sur un intervalle borné $\Omega =]0, L[$. Soit c la concentration du composant transporté, u la vitesse de Darcy qu'on suppose strictement positive, D le coefficient de diffusion et f le terme source, les conditions aux limites étant de type Dirichlet :

$$\begin{aligned} \frac{\partial c}{\partial t} - D \frac{\partial c}{\partial x^2} + u \frac{\partial c}{\partial x} &= f, \quad 0 < x < L, 0 < t < T \\ c(0, t) &= c_g(t), \quad c(L, t) = c_d(t), \\ c(x, 0) &= c_0(x) \quad \text{à } t = 0. \end{aligned} \quad (2.12)$$

Pour toutes les expériences numériques, nous avons utilisé un schéma aux différences finies avec $x_i = i \frac{L}{I}$, $i = 0, \dots, I$, et nous avons testé et comparé plusieurs schémas de discrétisation en temps.

Nous ne considérons que des schémas avec traitement implicite de la diffusion car dans le cas d'un traitement explicite de la diffusion la condition de stabilité impose que le pas de temps soit de l'ordre du carré du pas d'espace, ce qui rend le coût de calcul prohibitif.

Schéma totalement implicite en advection et en diffusion

Le schéma discret associé au problème (2.12) dont l'advection et la diffusion sont traitées, toutes les deux implicitement s'écrit :

$$\begin{aligned} \frac{c_i^{n+1} - c_i^n}{\Delta t} - D \frac{c_{i+1}^{n+1} - 2c_i^{n+1} + c_{i-1}^{n+1}}{\Delta x^2} + u \frac{c_i^{n+1} - c_{i-1}^{n+1}}{\Delta x} &= f_i^{n+1}, \quad \text{pour } i = 1, \dots, I-1 \\ c_0^{n+1} &= c_g(t^{n+1}), \quad c_I^{n+1} = c_d(t^{n+1}), \\ f_i^{n+1} &= f(x_i, t^{n+1}). \end{aligned} \quad (2.13)$$

et où le second membre vaut :

$$B_2 = \begin{pmatrix} (1 - a_v)c_1^n + a_v c_0^n + a_d c_0^{n+1} + f_1 \Delta t \\ (1 - a_v)c_2^n + a_v c_1^n + f_2 \Delta t \\ \vdots \\ (1 - a_v)c_{I-1}^n + a_v c_{I-2}^n + a_d c_{I-1}^{n+1} + f_{I-1} \Delta t \end{pmatrix}$$

Splitting : implicite en diffusion et explicite en advection

Soit Δt le pas de temps pour la diffusion. Il sera divisé en M pas de temps d'advection Δt_a tels que $\Delta t = M \Delta t_a$ avec $M \geq 1$, le pas de temps d'advection sera contrôlé par une condition CFL :

$$\Delta t_a \leq \frac{\Delta x}{u}.$$

Étape d'advection

L'intervalle $[t_n, t_{n+1}]$ est divisé en M intervalles $[t_{n,m}, t_{n,m+1}]$, $m = 0, \dots, M - 1$, avec $t_{n,0} = t_n$, $t_{n,M} = t_{n+1}$. On note par $c^{n,m}$ l'approximation de c à l'instant $t^{n,m}$ avec $c^{n,0} = c^n$, le schéma décentré explicite pour l'advection s'écrit :

$$\frac{c_i^{n,m+1} - c_i^{n,m}}{\Delta t_a} + u \frac{c_i^{n,m} - c_{i-1}^{n,m}}{\Delta x} = 0, \quad m = 0, \dots, M - 1, i = 0, \dots, I, \quad (2.16)$$

ou

$$c_i^{n,m+1} = c_i^{n,m} - \frac{\Delta t_a u}{\Delta x} (c_i^{n,m} - c_{i-1}^{n,m}), \quad m = 0, \dots, M - 1 \quad (2.17)$$

On calcule ainsi explicitement l'advection, et on obtient $c_i^{n,M}$ après M pas de temps d'advection. Cette concentration représente une prédiction de c^n obtenue en prenant en compte seulement l'advection.

Étape de diffusion

Étant donné $c_i^{n,M}$ obtenu dans l'étape d'advection, on calcule c_i^{n+1} par un schéma implicite, en résolvant l'équation de diffusion avec le pas de temps Δt

$$\frac{c_i^{n+1} - c_i^{n,M}}{\Delta t} - \frac{D}{\Delta x^2} (c_{i+1}^{n+1} - 2c_i^{n+1} + c_{i-1}^{n+1}) = f_i^{n+1} \quad (2.18)$$

ou

$$-a_d c_{i+1}^{n+1} + (1 + 2a_d) c_i^{n+1} - a_d c_{i-1}^{n+1} = c_i^{n,M} + \Delta t f_i^{n+1} \quad (2.19)$$

on obtient le système $A_3 X = B_3$ avec :

$$A_3 = \begin{pmatrix} 1 + 2c_d & -a_d & & & & & \\ -a_d & 1 + 2a_d & -a_d & & & & \\ & & \cdot & \cdot & & & \\ & & & & -a_d & 1 + 2a_d & -a_d \\ & & & & & -a_d & 1 + 2a_d \end{pmatrix}$$

et

$$B_3 = \begin{pmatrix} c_1^n + a_d c_0^{n+1} + f_1 \Delta t \\ c_2^n + f_2 \Delta t \\ \cdot \\ \cdot \\ c_{I-1}^n + a_d c_I^{n+1} + f_{I-1} \Delta t \end{pmatrix}$$

La solution analytique

On considère dans notre étude le transport d'un traceur dans un domaine monodimensionnel semi-infini, on impose une concentration constante en $x = 0$ pour $t \geq 0$, le modèle mathématique revient au problème décrit précédemment. On peut obtenir la solution analytique en utilisant la transformation de Laplace, et la solution exacte [84] s'écrit :

$$c(x, t) = \frac{c_0}{2} \left[\operatorname{erfc} \left(\frac{x - ut}{2\sqrt{Dt}} \right) + \exp \left(\frac{ux}{D} \right) \operatorname{erfc} \left(\frac{x + ut}{2\sqrt{Dt}} \right) \right]. \quad (2.20)$$

Comparaison des résultats numériques avec la solution analytique

Sur toutes les figures, les courbes rouges représentent les solutions analytiques et les courbes bleues représentent la concentration calculée. Dans tous les cas tests, on fixe la porosité 0.1, la diffusion 0.05 (sauf dans le cas d'advection pure, où on prend une diffusion égale à 10^{-22}), la vitesse de Darcy 1 (sauf dans le cas de diffusion pure, où on prend une vitesse égale à 0), on prend le pas d'espace $\Delta x = 0.05$, et on fait varier les pas de temps.

Diffusion pure

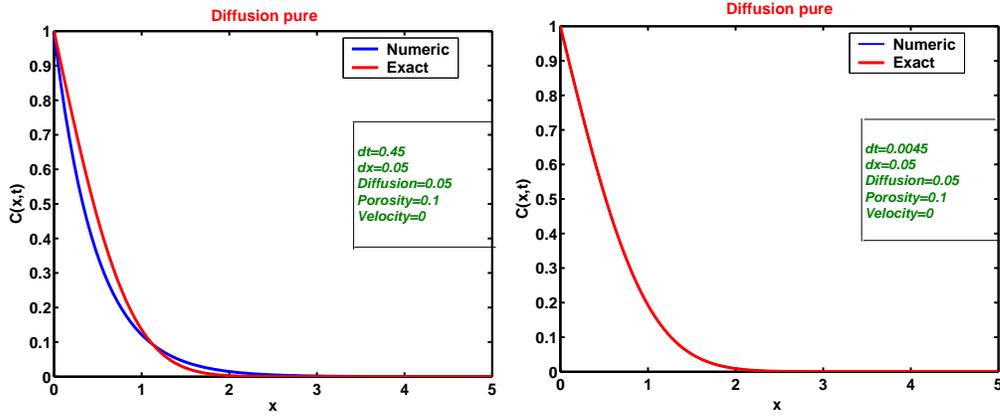


FIG. 2.1: Cas d'une diffusion pure : le pas de temps est 0.45 à gauche, et 0.0045 à droite.

Dans le cas de diffusion pure, les trois schémas donnent des résultats identiques car la diffusion est traitée de même manière par un schéma implicite en temps. Les figures (2.1) représentent les résultats obtenus pour une diffusion pure sans advection pour deux pas de temps différents. On constate l'effet du choix du pas de temps sur la solution calculée. L'erreur diminue avec le pas de temps.

Advection pure

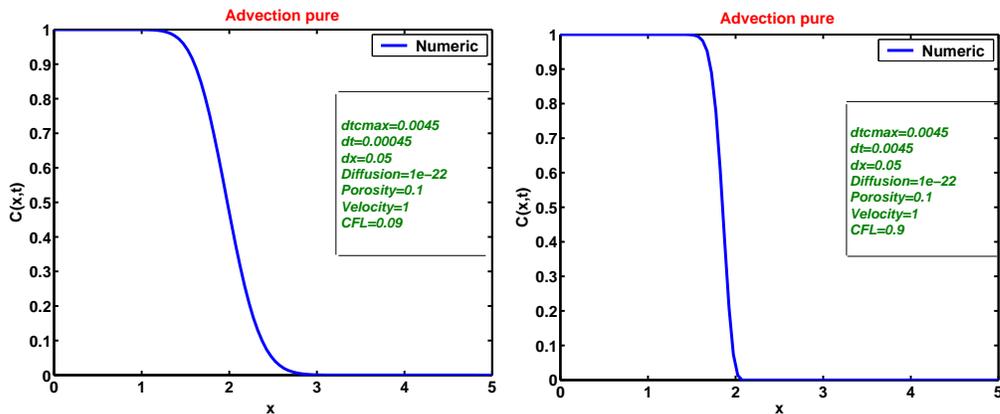


FIG. 2.2: Cas d'une advection pure : le pas de temps est 0.00045 à gauche, et 0.0045 à droite

Les figures (2.2) montrent l'influence du pas de temps sur la solution calculée pour le schéma explicite en temps. A droite la solution correspond à un pas de temps très proche du pas de temps maximum autorisé par la condition CFL. A gauche c'est

la solution avec un pas de temps dix fois plus petit. Pour ce schéma décentré amont, l'erreur diminue quand le pas de temps se rapproche du pas de temps maximum autorisé par la condition CFL.

Diffusion et advection

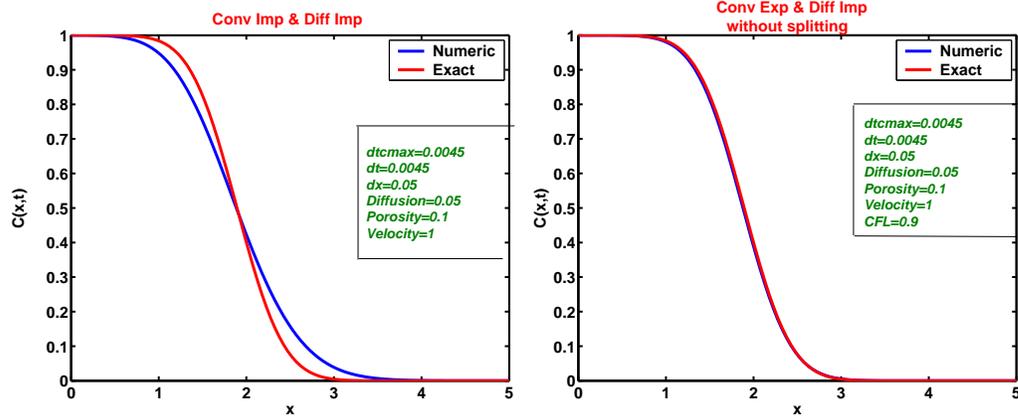


FIG. 2.3: Diffusion et advection sans splitting : à gauche diffusion et advection implicite, à droite diffusion implicite et advection explicite.

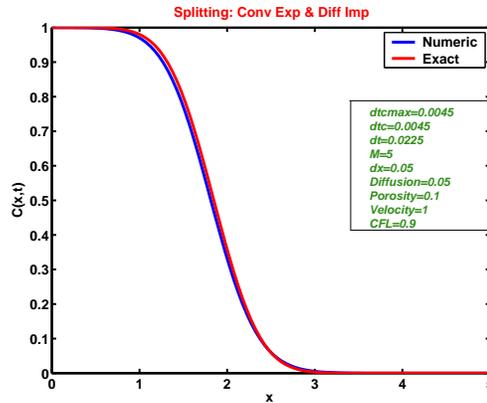


FIG. 2.4: Diffusion et advection avec splitting : le pas de temps de diffusion est dix fois plus grand que le pas de temps d'advection

Sur les figures (2.3), on compare le schéma (2.15) où l'advection est explicite (à droite) et le schéma (2.23) où l'advection est implicite (à gauche), la diffusion étant traitée implicitement dans les deux cas. Le pas de temps choisi est proche du pas de temps maximum autorisé par la condition CFL dans les deux cas. On constate

que, quand on traite implicitement l'advection, la diffusion numérique est plus importante.

La figure (2.4) montre un résultat obtenu avec décomposition des opérateurs de diffusion et d'advection (splitting). Ici le pas de temps d'advection est proche du pas de temps maximum autorisé par la condition CFL, et le pas de temps choisi pour la diffusion est cinq fois plus grand que le pas de temps d'advection, ce qui donne un rapport de cinq pas de temps d'advection pour un pas de temps de diffusion. On constate que la solution est aussi bonne que sur la figure (2.3) à droite mais pour un coût cinq fois inférieur.

D'après les résultats que nous avons obtenus, on constate qu'une méthode de décomposition des opérateurs de diffusion et d'advection, avec advection explicite, donne de bons résultats quand l'advection est importante. Le fait de séparer diffusion et advection nous permet d'utiliser un pas de temps optimal pour chacun des opérateurs. La séparation des opérateurs, pour les schémas que nous considérons, ne dégrade pas la qualité de la solution. Nous généralisons cette étude au cas des milieux hétérogènes, cf. section 5.2. Nous allons maintenant donner quelques éléments théoriques sur les schémas que nous avons présentés.

Ordres des schémas et erreurs d'approximation

Dans cette partie, on va étudier l'erreur d'approximation des trois schémas étudiés précédemment. Soit $c(x, t)$ la solution exacte (2.20). Pour calculer l'ordre du schéma, ou mesurer l'erreur d'une approximation numérique, on utilise des développements de Taylor successifs. On note h le pas d'espace et k le pas de temps. On a :

$$\begin{aligned}
 c(x, t \pm k) &= c(x, t) \pm k \frac{\partial c}{\partial t} + \frac{k^2}{2} \frac{\partial^2 c}{\partial t^2} + O(k^3), \\
 c(x \pm h, t) &= c(x, t) \pm h \frac{\partial c}{\partial x} + \frac{h^2}{2} \frac{\partial^2 c}{\partial x^2} \pm \frac{h^3}{6} \frac{\partial^3 c}{\partial x^3} + O(h^4), \\
 c(x - h, t - k) &= c(x, t) - h \frac{\partial c}{\partial x} - k \frac{\partial c}{\partial t} + \frac{h^2}{2} \frac{\partial^2 c}{\partial x^2} + \frac{k^2}{2} \frac{\partial^2 c}{\partial t^2} \\
 &\quad + hk \frac{\partial^2 c}{\partial x \partial t} + O(h^3) + O(k^3) + O(h^2k) + O(k^2h),
 \end{aligned} \tag{2.21}$$

Appliquant ces développements de Taylor aux combinaisons qui interviennent dans les schémas :

$$c_{i-1}^n - 2c_i^n + c_{i+1}^n = \frac{h^2}{2} \frac{\partial^2 c}{\partial x^2} + O(h^4),$$

$$c_i^{n-1} - c_{i-1}^{n-1} = h \frac{\partial c}{\partial x} - \frac{h^2}{2} \frac{\partial^2 c}{\partial x^2} - hk \frac{\partial^2 c}{\partial x \partial t} + O(h^3) + O(h^2k) + O(k^3) + O(k^2h),$$

Pour le schéma totalement implicite, on pose,

$$\epsilon_{im,im}^n = \frac{c_i^n - c_i^{n-1}}{k} - D \frac{c_{i+1}^n - 2c_i^n + c_{i-1}^n}{h^2} + u \frac{c_i^n - c_{i-1}^n}{h}, \quad (2.22)$$

on remplace ces termes par les formules de Taylor, et comme c est solution de l'équation (2.12), on obtient :

$$\epsilon_{im,im}^n = -\frac{k}{2} \frac{\partial^2 c}{\partial t^2} - \frac{h}{2} \frac{\partial^2 c}{\partial x^2} + O(h^2) + O(k^2). \quad (2.23)$$

Le schéma totalement implicite est donc un schéma consistant d'ordre 1 en espace et temps. De même façon, on calcule l'ordre de l'erreur produite par l'approximation explicite en advection et implicite en diffusion. On pose

$$\epsilon_{im,ex}^n = \frac{c_i^n - c_i^{n-1}}{k} - D \frac{c_{i+1}^n - 2c_i^n + c_{i-1}^n}{h^2} + u \frac{c_i^{n-1} - c_{i-1}^{n-1}}{h}, \quad (2.24)$$

On a :

$$\begin{aligned} c_i^{n-1} - c_{i-1}^{n-1} &= h \frac{\partial c}{\partial x} - hk \frac{\partial^2 c}{\partial x \partial t} - \frac{h^2}{2} \frac{\partial^2 c}{\partial x^2} + O(h^2) \\ \epsilon_{im,ex}^n &= \frac{\partial c}{\partial t} - D \frac{\partial c}{\partial x^2} + u \frac{\partial c}{\partial x} - uk \frac{\partial^2 c}{\partial x \partial t} - \frac{k}{2} \frac{\partial^2 c}{\partial t^2} - \frac{h}{2} \frac{\partial^2 c}{\partial x^2} + O(h^2) + O(k^2), \end{aligned} \quad (2.25)$$

et comme c est solution de l'équation (2.12), on obtient :

$$\epsilon_{im,ex}^n = -uk \frac{\partial^2 c}{\partial x \partial t} - \frac{k}{2} \frac{\partial^2 c}{\partial t^2} - \frac{h}{2} \frac{\partial^2 c}{\partial x^2} + O(h^2) + O(k^2), \quad (2.26)$$

ce qui montre que le schéma semi explicite est aussi d'ordre 1 en espace et en temps.

Pour montrer que l'erreur de séparation d'opérateurs à plusieurs pas de temps est d'ordre 1, on pose d'abord :

$$\epsilon_{sp,1}^n = \frac{c_i^{n+1} - c_i^{n,0}}{k} - D \frac{c_{i+1}^{n+1} - 2c_i^{n+1} + c_{i-1}^{n+1}}{h^2} + u \frac{c_i^{n,0} - c_{i-1}^{n,0}}{h},$$

qui correspond au schéma de séparation d'opérateurs avec un seul pas d'advection pour un pas de diffusion. Ce schéma est équivalent au schéma implicite en diffusion et explicite en advection sans décomposition d'opérateurs, on a montré précédemment qu'il est d'ordre 1 en espace et en temps. Et si on pose

$$\mathcal{A} = D \frac{\partial}{\partial x^2} \quad \text{et} \quad \mathcal{B} = -u \frac{\partial}{\partial x}$$

On a $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, alors l'erreur de séparations d'opérateurs est nulle dans ce cas, cf. section 4.

Avec une séparation d'opérateurs à plusieurs pas du temps d'advection, on a :

$$c_i^{n,M} = c_i^{n,0} - \frac{u\Delta t}{M\Delta x} \sum_{m=0}^{M-1} (c_i^{n,m} - c_{i-1}^{n,m}). \quad (2.27)$$

cette concentration est la somme des concentrations obtenues aux sous pas d'advection (2.17), soit

$$\frac{c_i^{n,0}}{k} = \frac{c_i^{n,M}}{k} + \frac{u}{Mh} \sum_{m=0}^{M-1} (c_i^{n,m} - c_{i-1}^{n,m}). \quad (2.28)$$

On peut écrire :

$$-\frac{c_i^{n,0}}{k} + u \frac{c_i^{n,0} - c_{i-1}^{n,0}}{h} = -\frac{c_i^{n,M}}{k} - \frac{u}{h} \left(\sum_{m=0}^{M-1} \frac{1}{M} (c_i^{n,m} - c_{i-1}^{n,m}) - (c_i^{n,0} - c_{i-1}^{n,0}) \right), \quad (2.29)$$

et maintenant on pose pour le schéma de séparation d'opérateurs avec plusieurs pas de temps d'advection :

$$\epsilon_{sp,M}^n = \frac{c_i^{n+1}}{k} - D \frac{c_{i+1}^{n+1} - 2c_i^{n+1} + c_{i-1}^{n+1}}{h^2} - \frac{c_i^{n,M}}{k} - \frac{u}{h} \left(\sum_{m=0}^{M-1} \frac{1}{M} (c_i^{n,m} - c_{i-1}^{n,m}) - (c_i^{n,0} - c_{i-1}^{n,0}) \right), \quad (2.30)$$

Cette dernière équation possède un terme $\epsilon = \frac{u}{h} \left(\sum_{m=0}^{M-1} \frac{1}{M} (c_i^{n,m} - c_{i-1}^{n,m}) - (c_i^{n,0} - c_{i-1}^{n,0}) \right)$ supplémentaire par rapport à l'équation de diffusion (2.18). Ce terme est dû à la division du pas de temps en plusieurs pas de temps d'advection, il est nul si $M=1$. Et comme pour chaque pas de temps d'advection, ce schéma est équivalent au schéma d'advection explicite qui est d'ordre 1 en espace et en temps, alors le schéma à plusieurs pas de temps d'advection est au plus d'ordre 1 en espace et en temps.

Après cette étude théorique de l'ordre, nous présentons une étude numérique pour comparer l'erreur d'approximation numérique avec la solution analytique pour les trois schémas.

Pour cela, on reprend les mêmes données des tests décrits dans la section 5.1, (paragraphe sur la comparaison des résultats d'advection-diffusion), sauf pour le cas splitting, pour le pas de temps de diffusion, on a pris le même pas de temps choisi pour les deux autres schémas, c-à-d le maximum autorisé par la condition CFL, et on a pris un pas de temps d'advection 5 fois plus petit que le pas de temps de diffusion, $M=5$. On considère différentes maillages $h = 1/10$, $h = 1/20$, $h = 1/40$, $h = 1/80$, $h = 1/160$, $h = 1/320$, pour calculer l'erreur en fonction de la discrétisation en espace. Pour les trois schémas, les conditions aux bords sont données exactement par la solution analytique.

Les erreurs correspondant à différentes maillages sont présentées sur la figure (FIG. 2.5).

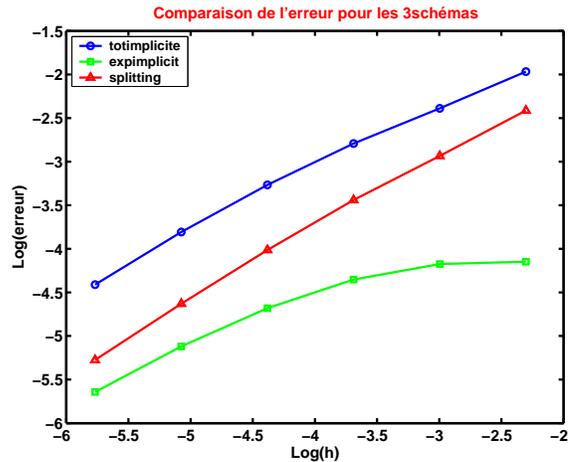


FIG. 2.5: Comparaison d'erreurs pour les trois schémas numériques avec la solution analytique

Pour les trois schémas, on constate que l'erreur diminue avec le pas d'espace h , les trois schémas sont d'ordre 1 en espace. Le splitting est "presque" aussi bon que le schéma semi explicite, mais il permet un plus grand pas de temps.

5.2 Étude du cas hétérogène

Dans cette partie, on résout l'équation d'advection-diffusion dans le cas hétérogène, on prend en compte l'hétérogénéité de la diffusion, de la porosité. Pour la discrétisation en espace, on va utiliser une méthode volumes finis sur un maillage irrégulier, et on va étudier et comparer trois schémas différents pour la discrétisation en temps.

Le modèle

On s'intéresse donc à l'étude de l'équation d'advection-diffusion dans le cas monodimensionnel sur un intervalle borné $\Omega =]0, L[$.

$$\phi \frac{\partial c}{\partial t} + \frac{\partial}{\partial x} \left(-D \frac{\partial c}{\partial x} + uc \right) = f, \quad 0 < x < L, 0 < t < T \quad (2.31)$$

Soit ϕ la porosité du milieu poreux et c la concentration d'un composant dissout dans l'eau, transporté par diffusion et advection. D le coefficient de diffusion, f le terme source et u la vitesse de l'eau supposé incompressible, ce qui implique en dimension 1 que u est constant et supposé positive.

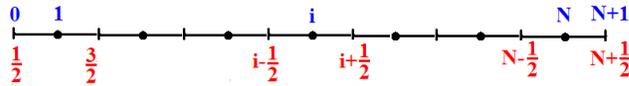
Conditions initiale et aux bords :

A l'équation (2.31), on associe une condition initiale $c(x, 0) = c_0(x)$ et des conditions aux bords. Dans la suite, on choisit d'imposer en $x = 0$ une condition de Dirichlet $c(0, t) = cd(t)$, et en $x = L$, on impose une condition de type Neumann c'est à dire qu'on impose un flux diffusif noté $\bar{\varphi}$ au dernier point, tel que $\int_{t_n}^{t_{n+1}} \varphi_{d, N+\frac{1}{2}} dt = \bar{\varphi}$.

a) Discrétisation en espace

On traite la discrétisation en espace et en temps séparément, on utilise différents schémas en temps pour les différents termes de l'opérateur du transport.

Pour la discrétisation en espace, et pour toutes les expériences numériques, on utilise une méthode de volume finis centrés sur les mailles. Les mailles sont notées par les indices $i = 1, \dots, N$ et les extrémités des mailles par les indices $i + 1/2$, $i = 0, \dots, N$. L'intervalle $[0, L]$ est divisé en N intervalles $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ de longueur h_i , $i = 1, \dots, N$ avec $x_{\frac{1}{2}} = 0$, $x_{N+\frac{1}{2}} = L$. La concentration c est approchée par une fonction constante par maille et les valeurs par maille sont notées c_i , $i = 1, \dots, N$ plus une valeur à chacune des deux extrémités de l'intervalle $]0, L[$ notées c_0 et c_{N+1} .



On écrit (2.31) sous la forme :

$$\phi \frac{\partial c}{\partial t} + \frac{\partial \varphi}{\partial x} = f, \quad (2.32)$$

où : $\varphi(x, t) = -D \frac{\partial c}{\partial x} + uc$ est le flux de soluté traversant le point x à l'instant t . Ce

flux est la somme de deux flux $\varphi = \varphi_d + \varphi_a$, un flux de diffusion $\varphi_d = -D \frac{\partial c}{\partial x}$ et un flux d'advection $\varphi_a = uc$.

On intègre l'équation (2.32) sur $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, on obtient :

$$\phi_i h_i \frac{dc_i}{dt} + \varphi_{i+\frac{1}{2}} - \varphi_{i-\frac{1}{2}} = f_i h_i,$$

$$\phi_i h_i \frac{dc_i}{dt} + \varphi_{d, i+\frac{1}{2}} + \varphi_{a, i+\frac{1}{2}} - \varphi_{d, i-\frac{1}{2}} - \varphi_{a, i-\frac{1}{2}} = f_i h_i$$

où les flux $\varphi_{d, i+\frac{1}{2}}$ et $\varphi_{a, i+\frac{1}{2}}$ sont définis par :

$$\varphi_{d, i+\frac{1}{2}} = -D_{i+\frac{1}{2}} \left(\frac{c_{i+1} - c_i}{h_{i+\frac{1}{2}}} \right), \quad \varphi_{a, i+\frac{1}{2}} = uc_i.$$

On suppose que ϕ_i et D_i , la porosité (resp. la diffusion) sont constantes sur la maille i . On choisit d'évaluer $D_{i+\frac{1}{2}}$ en faisant une moyenne harmonique de D_i et D_{i+1} et on choisit d'évaluer le flux advectif de la manière à ce qu'il soit décentré en amont. La moyenne harmonique est l'inverse de la moyenne des inverses.

$$D_{i+\frac{1}{2}} = \frac{2D_i D_{i+1}}{D_i + D_{i+1}}, \quad \text{avec} \quad D_{\frac{1}{2}} = D_1 \text{ et } D_{N+\frac{1}{2}} = D_N, \quad \text{et} \quad h_{i+\frac{1}{2}} = \frac{h_i + h_{i+1}}{2}.$$

b) Discrétisation en temps

Comme dans la section précédente l'opérateur du transport contient un terme diffusif et un autre advectif, et on utilise différentes discrétisation en temps pour les différents termes.

Advection et diffusion : schéma implicite

On discrétise l'intervalle $[0, T]$ par le pas de temps Δt , et on note par c_i^n la valeur de $c_i(n\Delta t)$. Le schéma totalement implicite s'écrit :

$$\phi_i h_i (c_i^{n+1} - c_i^n) + (\varphi_{d,i+\frac{1}{2}}^{n+1} + \varphi_{a,i+\frac{1}{2}}^{n+1} - \varphi_{d,i-\frac{1}{2}}^{n+1} - \varphi_{a,i-\frac{1}{2}}^{n+1}) \Delta t = f_i h_i \Delta t.$$

On a alors après remplacement des quantités :

$$\phi_i h_i c_i^{n+1} - D_{i+\frac{1}{2}} \Delta t \left(\frac{c_{i+1}^{n+1} - c_i^{n+1}}{h_{i+\frac{1}{2}}} \right) + D_{i-\frac{1}{2}} \Delta t \left(\frac{c_i^{n+1} - c_{i-1}^{n+1}}{h_{i-\frac{1}{2}}} \right) + u \Delta t c_i^{n+1} - u \Delta t c_{i-1}^{n+1} = \phi_i h_i c_i^n + f_i h_i \Delta t$$

En regroupant les termes on obtient :

$$\left(-\frac{D_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}} \Delta t - u \Delta t \right) c_{i-1}^{n+1} + \left(\phi_i h_i + \frac{D_{i+\frac{1}{2}}}{h_{i+\frac{1}{2}}} \Delta t + \frac{D_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}} \Delta t + u \Delta t \right) c_i^{n+1} - \left(\frac{D_{i+\frac{1}{2}}}{h_{i+\frac{1}{2}}} \Delta t \right) c_{i+1}^{n+1} = \phi_i h_i c_i^n + f_i h_i \Delta t.$$

Condition aux bords :

Pour $i=1$, on a imposé une condition de Dirichlet à gauche donc on connaît la concentration au point x_0

$$\int_{x_{\frac{1}{2}}}^{x_{\frac{3}{2}}} \phi_1 (c_1^{n+1} - c_1^n) dx + \int_{t_n}^{t_{n+1}} (\varphi_{d,\frac{3}{2}} + \varphi_{a,i+\frac{3}{2}} - \varphi_{d,\frac{1}{2}} - \varphi_{a,\frac{1}{2}}) dt = \int_{t_n}^{t_{n+1}} f_1 h_1 dt,$$

soit

$$\left(\phi_1 h_1 + \frac{D_{\frac{3}{2}}}{h_{\frac{3}{2}}} \Delta t + u \Delta t + \frac{D_{\frac{1}{2}}}{h_{\frac{1}{2}}} \Delta t \right) c_1^{n+1} - \frac{D_{\frac{3}{2}}}{h_{\frac{3}{2}}} \Delta t c_2^{n+1} = \phi_1 h_1 c_1^n + \left(\frac{D_{\frac{1}{2}}}{h_{\frac{1}{2}}} \Delta t + \Delta t u \right) c_d^{n+1} + f_1 h_1 \Delta t$$

Pour $i=N$, on a

$$\int_{x_{N-\frac{1}{2}}}^{x_{N+\frac{1}{2}}} \phi_N (c_N^{n+1} - c_N^n) dx + \int_{t_n}^{t_{n+1}} (\varphi_{d,N+\frac{1}{2}} + \varphi_{a,N+\frac{1}{2}} - \varphi_{d,N-\frac{1}{2}} - \varphi_{a,N-\frac{1}{2}}) dt = \int_{t_n}^{t_{n+1}} f_N h_N dt$$

Or par la condition de flux diffusif imposée dans les hypothèses on a :

$$\phi_N h_N (c_N^{n+1} - c_N^n) + (\bar{\varphi} + u c_N^{n+1} + \frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}} (c_N^{n+1} - c_{N-1}^{n+1}) - u c_{N-1}^{n+1}) \Delta t = f_N h_N \Delta t,$$

soit

$$c_{N-1}^{n+1} \left(-\frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}} - u \right) \Delta t + c_N^{n+1} \left(\phi_N h_N + u \Delta t + \frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}} \Delta t} \right) = \phi_N h_N c_N^n - \bar{\varphi} \Delta t + f_N h_N \Delta t.$$

Schéma implicite en diffusion et explicite en advection

Cette fois ci, on veut traiter l'advection explicitement. Le schéma implicite en diffusion et explicite en advection s'écrit :

$$\phi_i h_i (c_i^{n+1} - c_i^n) + (\varphi_{d,i+\frac{1}{2}}^{n+1} + \varphi_{a,i+\frac{1}{2}}^n - \varphi_{d,i-\frac{1}{2}}^{n+1} - \varphi_{a,i-\frac{1}{2}}^n) \Delta t = f_i h_i \Delta t$$

En remplaçant et regroupant les termes on a :

$$\begin{aligned} -\left(\frac{D_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}} \Delta t \right) c_{i-1}^{n+1} + \left(\phi_i h_i + \frac{D_{i+\frac{1}{2}}}{h_{i+\frac{1}{2}}} \Delta t + \frac{D_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}} \Delta t \right) c_i^{n+1} - \frac{D_{i+\frac{1}{2}}}{h_{i+\frac{1}{2}}} \Delta t c_{i+1}^{n+1} = \\ (\phi_i h_i - u \Delta t) c_i^n + u \Delta t c_{i-1}^n + f_i h_i \Delta t. \end{aligned}$$

Condition aux bords :

Pour $i=1$, on a imposé une condition de Dirichlet à gauche

$$\phi_1 h_1 (c_1^{n+1} - c_1^n) + \left(\frac{-D_{\frac{3}{2}}}{h_{\frac{3}{2}}} c_2^{n+1} + \frac{D_{\frac{3}{2}}}{h_{\frac{3}{2}}} c_1^{n+1} + u c_1^n + \frac{D_{\frac{1}{2}}}{h_{\frac{1}{2}}} c_1^{n+1} - \frac{D_{\frac{1}{2}}}{h_{\frac{1}{2}}} c_d^{n+1} - u c_d^n \right) \Delta t = f_1 h_1 \Delta t.$$

On obtient donc finalement comme expression :

$$\begin{aligned} \left(\phi_1 h_1 + \frac{D_{\frac{3}{2}}}{h_{\frac{3}{2}}} \Delta t + \frac{D_{\frac{1}{2}}}{h_{\frac{1}{2}}} \right) c_1^{n+1} - \frac{D_{\frac{3}{2}}}{h_{\frac{3}{2}}} \Delta t c_2^{n+1} = \\ (\phi_1 h_1 - u \Delta t) c_1^n + (u \Delta t) c_d^n + \frac{D_{\frac{1}{2}}}{h_{\frac{1}{2}}} \Delta t c_d^{n+1} + f_1 h_1 \Delta t. \end{aligned}$$

Pour $i=N$, on a

$$\int_{x_{N-\frac{1}{2}}}^{x_{N+\frac{1}{2}}} \phi_N (c_N^{n+1} - c_N^n) dx + \int_{t_n}^{t_{n+1}} (\varphi_{d,N+\frac{1}{2}} + \varphi_{a,N+\frac{1}{2}} - \varphi_{d,N-\frac{1}{2}} - \varphi_{a,N-\frac{1}{2}}) dt = \int_{t_n}^{t_{n+1}} f_N h_N dt.$$

Or par la condition de flux diffusif imposée, on a :

$$\phi_N h_N (c_N^{n+1} - c_N^n) + (\bar{\varphi} + u c_N^n + \frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}} (c_N^{n+1} - c_{N-1}^{n+1}) - u c_{N-1}^n) \Delta t = f_N h_N \Delta t.$$

$$c_{N-1}^{n+1} (-\frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}}) \Delta t + c_N^{n+1} (\phi_N h_N + \frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}}) \Delta t = \phi_N h_N c_N^n + (u c_{N-1}^n - u c_N^n - \bar{\varphi}) \Delta t + f_N h_N \Delta t.$$

Splitting : Méthode de décomposition d'opérateurs

Maintenant, on veut utiliser la méthode de décomposition d'opérateurs. Dans un premier temps, on calcule la concentration $c_i^{n,M}$ utilisant M pas d'advection, ensuite on résout l'étape de diffusion en démarant avec cette concentration, $c_i^{n,M}$.

Étape d'advection

On s'intéresse dans cette étape à l'équation suivante : $\phi \frac{\partial c}{\partial t} + \frac{\partial \varphi_a}{\partial x} = 0$.
L'intervalle $[t_n, t_{n+1}]$ est divisé en M intervalles $[t_{n,m}, t_{n,m+1}]$, $m=0, \dots, M-1$, tel que $t_{n,0} = t_n$, $t_{n,M} = t_{n+1}$, $\Delta t = M \Delta t_a$ et Δt_a est contrôlé par la condition CFL.

$$\Delta t_a \leq \frac{\phi_i h_i}{u}.$$

On appelle $c_i^{n,m}$ l'approximation de la concentration c à l'instant $t_{n,m}$ avec $c^{n,0} = c^n$.
Le schéma d'Euler explicite nous donne pour l'équation d'advection :

$$\phi_i \left(\frac{c_i^{n,m+1} - c_i^{n,m}}{\Delta t_a} \right) + \frac{1}{h_i} (\varphi_{a,i+\frac{1}{2}}^{n,m} - \varphi_{a,i-\frac{1}{2}}^{n,m}) = 0, \quad i = 1, \dots, N, m = 0, \dots, M-1 \quad (2.33)$$

ou $\varphi_{a,i+\frac{1}{2}}^{n,m} = u c_i^{n,m}$.

On obtient à la fin de ces M pas de temps d'advection une quantité $c^{n,M}$ qui représente en fait une prédiction de la concentration obtenue en prenant en compte que l'advection.

Étape de diffusion

On s'intéresse maintenant à l'équation :

$$\phi \frac{\partial c}{\partial t} + \frac{\partial \varphi_d}{\partial x} = f. \quad (2.34)$$

On utilise un schéma implicite en temps, on intègre l'équation (2.34) sur $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], [t_n, t_{n+1}]$ et on obtient

$$\phi_i(c_i^{n+1} - c_i^n)h_i - (\varphi_{d,i+\frac{1}{2}}^{n+1} - \varphi_{d,i-\frac{1}{2}}^{n+1})\Delta t = 0.$$

En remplaçant et regroupant les termes après intégration on obtient :

$$\left(-\frac{D_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}}\Delta t\right)c_{i-1}^{n+1} + \left(\phi_i h_i + \frac{D_{i+\frac{1}{2}}}{h_{i+\frac{1}{2}}}\Delta t + \frac{D_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}}\Delta t\right)c_i^{n+1} - \frac{D_{i+\frac{1}{2}}}{h_{i+\frac{1}{2}}}\Delta t c_{i+1}^{n+1} = \phi_i h_i c_i^n f_i h_i \Delta t.$$

Condition aux bords :

Pour $i=1$, on a imposé comme avant une condition de dirichlet sur ce bord.
 Pour $i=N$, De la même manière qu'auparavant on impose une condition de flux diffusif, on obtient donc :

$$c_{N-1}^{n+1}\left(-\frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}}\Delta t\right) + c_N^{n+1}\left(\phi_N h_N + \frac{D_{N-\frac{1}{2}}}{h_{N-\frac{1}{2}}}\Delta t\right) = \phi_N h_N c_N^n - \bar{\varphi}\Delta t$$

Résultats numériques

Cas test : milieu hétérogène

Pour ce cas test, on prend un milieu hétérogène, composé de deux milieux A et B , le milieu A se caractérise par une faible porosité et diffusion par rapport au milieu B. Le point de discontinuité entre les deux milieux est le point $x = 2$.

	Milieu A	Milieu B
Porosité	1	2
Diffusion	10^{-2}	1
Longueur	2	3
Vitesse	10	10

On veut comparer les résultats données par les trois schémas pour une résolution du problème de diffusion et advection sur ce cas test hétérogène. On prend comme pas de temps $\Delta t = 9.10^{-4}$, et $\Delta tc = 9.10^{-5}$ le pas de temps d'advection pour le schéma splitting avec $M = 10$.

L'effet du maillage

Dans un premier test, on impose une condition de dirichlet au bord gauche $cd = 1$, et on considère une condition initiale nulle.

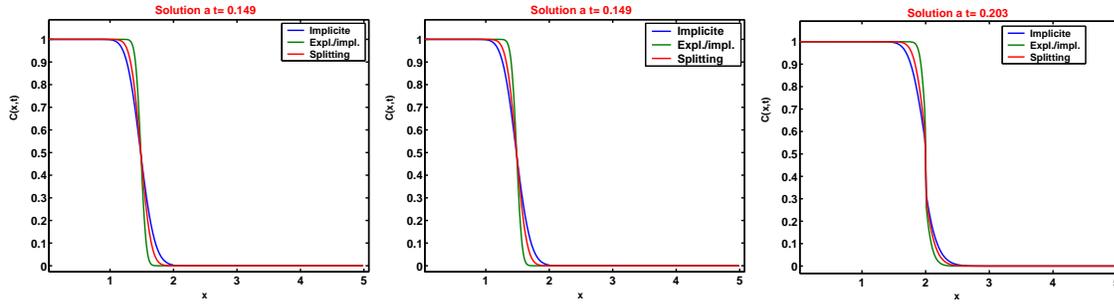


FIG. 2.6: Évolution de la concentration (maillage : nptsA=200 et nptsB = 150)

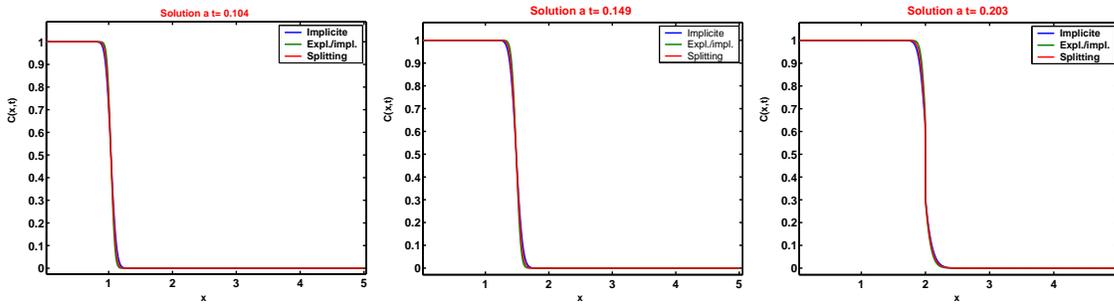


FIG. 2.7: Evolution de la concentration (maillage : nptsA = 1600 et nptsB = 1200)

Les trois premières figures (2.6) représentent l'évolution de la concentration calculée par les trois schémas à trois instants différents pour un maillage grossier, on note par nptsA le nombre de points du domaine A et nptsB le nombre du points du domaine B (nptsA=200 et nptsB = 150). Les figures (2.7) représentent les résultats pour un maillage plus fin (nptsA = 1600 et nptsB = 1200). Le maillage grossier n'est pas assez fin, on voit des différences.

Dans un deuxième test, on considère la condition initiale $c_0 = \exp[-3(1/2-x)^2]$, avec des conditions Dirichlet nulles sur les bords.

Les trois premières figures (2.8) représentent l'évolution de la concentration pour le maillage grossier, et les figures (2.9) représentent les résultats pour un maillage plus fin.

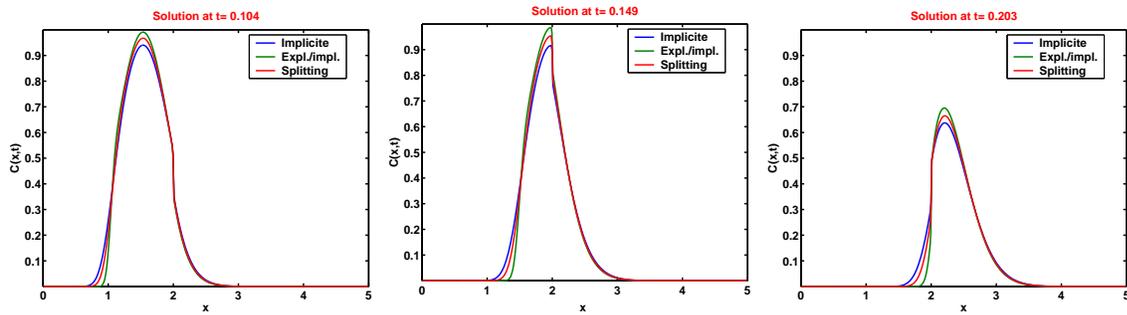


FIG. 2.8: Évolution de la concentration avec un maillage grossier (maillage : $nptsA=200$ et $nptsB = 150$)

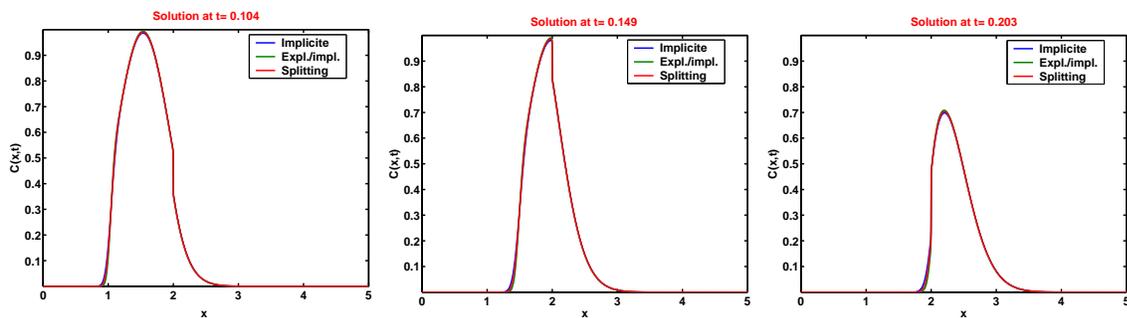


FIG. 2.9: Évolution de la concentration avec un maillage raffiné (maillage : $nptsA = 1600$ et $nptsB = 1200$)

On constate que la méthode splitting pour le transport donne des résultats de bonne qualité aussi bien pour le cas homogène que pour le cas hétérogène, et cette qualité ne se dégrade pas en raffinant le maillage.

5.3 Etude de convergence numérique

Dans cette partie, on étudie la convergence numérique des 3 schémas pour le cas hétérogène.

On reprend les mêmes données physiques et initiales que celles du deuxième test, cas hétérogène. Pour la discrétisation spatiale, on considère toujours que le pas d'espace h_B du domaine B est deux fois plus grand que le pas d'espace h_A du domaine A.

On prend un maillage très fin $h_A = 1/2^{10}$ dont la solution numérique représentera la solution exacte, et on prend une suite de maillage grossiers $h_A = \{1/2^q\}_{q=1, \dots, 9}$ et

on compare la différence.

Pour les schémas implicite et semi explicite, on prend comme un pas de temps $\Delta t = \Delta t_{\max}$ le maximum autorisé par la condition CFL, et pour le schéma splitting, le pas de temps d'advection $\Delta t_c = \Delta t_{\max}$ et le pas de temps de diffusion est M fois plus grand que le pas de temps d'advection $\Delta t = M\Delta t_{\max}$, avec $M=5$ ou $M=1$.

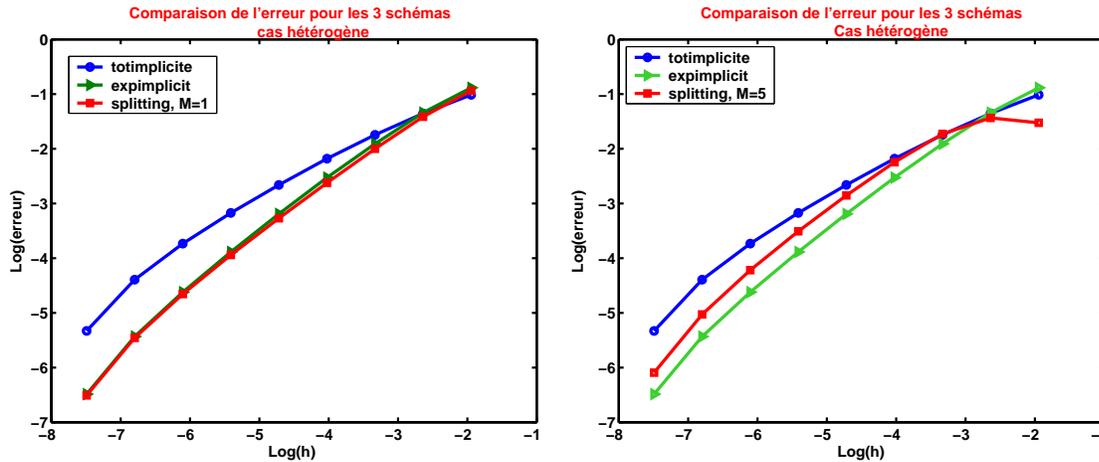


FIG. 2.10: Comparaison d'erreurs pour les 3 schémas avec une solution fine

Pour les 3 schémas, on constate que l'erreur diminue avec le pas d'espace h . Le schéma splitting est aussi bon que le schéma semi explicite, pour un coût M fois plus faible.

6 Mise en oeuvre du transport par décomposition d'opérateurs en 3D dans le logiciel 3FLO

ITASCA Consultants développe depuis plus de dix ans un simulateur hydrogéologique qui incorpore un module géochimique performant. Ce logiciel, nommé 3FLO, est notamment utilisé dans le cadre de coopérations internationales pour des projets de recherches et d'ingénierie appliquées au stockage de déchets radioactifs.

Le logiciel 3FLO inclut un couplage explicite simple entre le module de transport de soluté, basé sur la méthode des particules mobiles avec marche aléatoire, et le module de géochimie [10], [69]. Cette approche a permis à ITASCA Consultants de réaliser des simulations complexes intéressantes, mais elle a aussi montré ses limites en terme de précision et de vitesse de calcul.

La méthode des particules mobiles avec marche aléatoire montre certaines limites, essentiellement deux :

- Le nombre très élevé de particules qu’il est parfois nécessaire d’injecter dans le modèle pour obtenir une précision acceptable ;
- La précision obtenue par cette méthode de transport est a priori insuffisante pour envisager un couplage autre qu’explicite qui, pour être acceptable, oblige à une contrainte trop forte sur le pas de temps.

Ma contribution du développement de ce code concerne uniquement le transport sans toucher le code de la géochimie.

Ce qui existe déjà dans la version originale de 3Flo :

Pour l’équation d’écoulement :

Écoulement stationnaire : résolution par Galerkin ou Mixte-Hybride,

Écoulement transitoire : uniquement par Galerkin,

Pour l’équation du transport : la méthode des particules.

Pour mettre en oeuvre le transport par la décomposition d’opérateurs dans 3FLO, il me fallait modifier le code pour pouvoir simuler l’écoulement transitoire par la méthode mixte-hybride, puis adapter ce développement pour la résolution de la partie diffusive du transport car le problème de diffusion pure est similaire à un problème d’écoulement transitoire, ensuite implémenter la partie advective par la méthode volume finis pour pouvoir simuler le transport.

6.1 L’écoulement transitoire par la méthode mixte hybride

On veut utiliser une méthode mixte hybride en s’inspirant du travail de Chavent et Roberts [21]. Pour cette formulation, la loi de Darcy et l’équation de conservation de la masse sont discrétisées simultanément pour donner une approximation de \vec{u} et p .

Résolution

Soit Ω un ouvert de \mathbb{R}^3 , et \mathcal{T}_h une discrétisation en tétraèdres de Ω , où h est le paramètre du maillage indiquant le diamètre maximal d’une maille $K \in \mathcal{T}_h$. Soit $RT_0(K)$ l’espace de Raviart-Thomas de plus bas degré sur chaque élément K , et w_F une fonction de base de l’espace $RT_0(K)$ pour la face F du bord ∂K .

Pour chaque élément K de \mathcal{T}_h , w_F est défini par un polynome de la forme :

$$w_F(\vec{X}) = \vec{a} + b\vec{X}, \quad \vec{a} \in \mathbb{R}^3, b \in \mathbb{R}, \quad \forall \vec{X} \in K.$$

Les fonctions w_F satisfont les propriétés suivantes :

- Pour chaque $K \in \mathcal{T}_h$ et pour chaque $F \subset \partial K$,
- $\nabla \cdot w_F$ est constant sur l'élément K et $\int_K \nabla \cdot w_F = 1$.
 - $\nu_F \cdot w_{F'} = \frac{\delta_{F,F'}}{|F'|} \quad \forall F' \subset \partial K$.

où ν_F est le vecteur normal unitaire à F , dirigé vers l'extérieur de l'élément K , et $|F'|$ est la mesure de F' . La vitesse $\vec{u}_K \in RT_0(K)$ sur chaque élément K , est représentée par une combinaison des fonctions de bases w_F .

$$\vec{u}_K = \sum_{F \subset \partial K} Q_{K,F} w_F. \quad (2.35)$$

la vitesse \vec{u}_K a les propriétés suivantes :

- $\nabla \cdot \vec{u}_K$ est constant sur l'élément K .
- $\int_F \nu_{K,F} \cdot \vec{u}_K = Q_{K,F}$.

Par la connaissance des flux $Q_{K,F}$ à travers les arêtes F de K , on peut calculer la vitesse \vec{u} en tout point dans l'élément K .

Discrétisation de la loi de Darcy

Dans cette partie, on va écrire une forme discrète de l'équation d'écoulement (2.1). En inversant le tenseur de perméabilité dans cette équation, on obtient :

$$\mathcal{K}^{-1} \vec{u} = -\nabla p,$$

Soit $\vec{s}_K \in RT_0(K)$ une fonction test, on multiplie cette équation par la fonction \vec{s}_K et on intègre sur l'élément K . Par utilisation de la formule de Green, on obtient :

$$\int_K (\mathcal{K}^{-1} \vec{u}) \cdot \vec{s}_K = \int_K p \nabla \cdot \vec{s}_K - \int_{\partial K} p \vec{s}_K \cdot \vec{\nu} \quad \forall \vec{s}_K \in RT_0(K), \quad (2.36)$$

On approche $\vec{u}|_K$ par $\vec{u}_K \in RT_0(K)$, et $p|_K$ par une fonction constante $p_K \in \mathbb{R}$ et on approche indépendamment $p|_K$ sur la facette F du bord ∂K par $tp_{K,F} \in \mathbb{R}$. Soit \mathcal{K}_K la perméabilité, supposée constante sur l'élément K .

L'équation (2.36) devient :

$$\int_K \mathcal{K}_K^{-1} \vec{u}_K \cdot \vec{s}_K = \int_K p_K \nabla \cdot \vec{s}_K - \sum_{F \subset \partial K} \int_F tp_{K,F} \vec{s}_K \cdot \vec{\nu}_{K,F} \quad \forall \vec{s} \in RT_0(K). \quad (2.37)$$

On remplace successivement la fonction test \vec{s}_K par les fonctions de base $\vec{w}_{F'}$, $F' \subset \partial K$. Comme p_K est constante dans K et $tp_{K,F}$ est constante sur la face F , on peut écrire :

$$\sum_{F \subset \partial K} Q_{K,F} \int_K (\mathcal{K}_K^{-1} \vec{w}_F) \cdot \vec{w}_{F'} = p_K \int_K \nabla \cdot \vec{w}_{F'} - \sum_{F \subset \partial K} tp_{K,F} \int_F \vec{w}_{F'} \cdot \vec{\nu}_{K,F}. \quad (2.38)$$

En prenant en compte les propriétés des fonctions de bases, on obtient :

$$\sum_{F' \subset \partial K} Q_{K,F} (M_K)_{F',F} = p_K - tp_{K,F'} \quad \forall F' \subset \partial K, \quad (2.39)$$

où

$$(M_K)_{F',F} = \int_K (\mathcal{K}_K^{-1} \vec{w}_F) \cdot \vec{w}_{F'}. \quad (2.40)$$

L'écriture matricielle de ces équations donne le système suivant :

$$M_K Q_K = p_K e_K - tp_K, \quad (2.41)$$

où M_K est la matrice élémentaire et Q_K , tp_K et e_K sont des vecteurs locaux tels que :

$$M_K = ((M_K)_{F',F})_{F',F \subset \partial K},$$

$$Q_K = \{Q_{K,F}\}_{F \subset \partial K}, \quad tp_K = \{tp_{K,F}\}_{F \subset \partial K} \quad \text{et} \quad e_K = \{1\}_{F \subset \partial K}.$$

La matrice M_K est symétrique définie positive, donc inversible.

$$Q_K = M_K^{-1} (p_K e_K - tp_K) \quad (2.42)$$

Le flux à travers la face s'exprime par :

$$Q_{K,F} = \alpha_{K,F} p_K - \sum_{F' \subset \partial K} (M_K^{-1})_{F',F} tp_{K,F'}, \quad (2.43)$$

$$\text{où : } \alpha_{K,F} = \sum_{F' \subset \partial K} (M_K^{-1})_{F',F}.$$

Pour chaque face F qui n'appartient pas au bord de Ω , on a deux valeurs de la pression $tp_{K,F}$ et $tp_{K',F}$ et deux valeurs de flux $Q_{K,F}$ et $Q_{K',F}$. Pour avoir une approximation consistante de p et \vec{u} sur Ω , on ajoute deux équations qui expriment la continuité de la pression et le flux :

$$tp_{K,F} = tp_{K',F},$$

et

$$Q_{K,F} + Q_{K',F} = 0.$$

Discrétisation de l'équation de conservation

Maintenant, on discrétise l'équation de conservation de la masse :

$$S \frac{\partial p}{\partial t} - \nabla \cdot \vec{u} = f. \quad (2.44)$$

On multiplie l'équation (2.44) par une fonction r dans $L^2(K)$ et on l'intègre sur K , on obtient :

$$\int_K S \frac{\partial p}{\partial t} r + \int_K \nabla \cdot \vec{u} r = \int_K f r. \quad (2.45)$$

On note S_K , p_K des approximations de S resp p , constantes sur K et on prend $r = 1_K$ la fonction caractéristique de K . En utilisant le schéma d'Euler implicite pour la discrétisation en temps, on obtient :

$$\int_K S_K \frac{p_K^{n+1} - p_K^n}{\Delta t} 1_K + \int_K \nabla \cdot \vec{u} 1_K = \int_K f 1_K, \quad (2.46)$$

et la vitesse élémentaire \vec{u}_K satisfait :

$$\nabla \cdot \vec{u}_K = \frac{1}{|K|} \sum_{F \subset \partial K} Q_{K,F} = \text{constante sur } K.$$

Soit $\mathcal{F}_K = \frac{1}{|K|} \int_K f 1_K$, alors

$$S_K |K| \frac{p_K^{n+1} - p_K^n}{\Delta t} + e_K^T Q_K = \mathcal{F}_K. \quad (2.47)$$

Finalement, on obtient l'équation de bilan suivante :

$$S_K |K| \frac{p_K^{n+1}}{\Delta t} + e_K^T Q_K = S_K |K| \frac{p_K^n}{\Delta t} + \mathcal{F}_K, \quad (2.48)$$

ou encore

$$S_K |K| \frac{p_K^{n+1}}{\Delta t} + \sum_{F \subset \partial K} Q_{K,F} = S_K |K| \frac{p_K^n}{\Delta t} + \mathcal{F}_K, \quad (2.49)$$

On peut calculer les pressions moyenne p_K et les flux $Q_{K,F}$ localement une fois connaître les traces de la pression $tp_{K,F}$, et cela en utilisant les deux équations (2.42) et (2.48) et résolvant le système :

$$\begin{pmatrix} \frac{S_K |K|}{\Delta t} & e_K^T \\ -e_K & M_K \end{pmatrix} \begin{pmatrix} p_K^{n+1} \\ Q_K \end{pmatrix} = \begin{pmatrix} S_K |K| \frac{p_K^n}{\Delta t} + \mathcal{F}_K \\ -tp_K \end{pmatrix} \quad (2.50)$$

Par hybridation des équations et assemblage, on peut écrire le système globale dont les inconnus sont uniquement les traces de la pression $tp_{K,F}$.

Hybridization

Pour l'hybridization des équations, on élimine les pressions moyennes dans les éléments et les flux pour avoir un système dont les inconnues sont uniquement les traces de la pression, en suivant les étapes suivantes :

- 1) On remplace les flux de l'équation de conservation (2.49) par leur expression obtenue par l'équation (2.43), on obtient

$$S_K|K|\frac{p_K^{n+1}}{\Delta t} + \sum_{F \subset \partial K} \left[\alpha_{K,F} p_K^{n+1} - \sum_{F' \subset \partial K} (M_K^{-1})_{F',F} t p_{K,F'}^{n+1} \right] = S_K|K|\frac{p_K^n}{\Delta t} + \mathcal{F}_K,$$

$$S_K|K|\frac{p_K^{n+1}}{\Delta t} + \alpha_K p_K^{n+1} - \sum_{F' \subset \partial K} \alpha_{K,F'} t p_{K,F'}^{n+1} = S_K|K|\frac{p_K^n}{\Delta t} + \mathcal{F}_K,$$

où $\alpha_K = \sum_{F \subset \partial K} \alpha_{K,F}$.

- 2) On exprime la pression moyenne dans l'élément en fonction des traces de la pression à partir de l'équation de conservation.

Soit $\lambda_K = \frac{\alpha_K \Delta t}{S_K|K|}$ et $\beta_K = \frac{\lambda_K}{1 + \lambda_K}$,

$$p_K^{n+1} = (1 - \beta_K) p_K^n + \frac{\beta_K}{\alpha_K} F_K + \beta_K \sum_{F' \subset \partial K} \frac{\alpha_{K,F'}}{\alpha_K} t p_{K,F'}^{n+1},$$

- 3) l'équation en trace de la pression est calculée en remplaçant la pression moyenne par son expression obtenue à l'étape 2 dans l'équation décrivant la continuité de flux.

On a pour toute face F : $Q_{K,F}^{n+1} + Q_{K',F}^{n+1} = 0$.

$$\alpha_{K,F} p_K + \alpha_{K,F} p_{K'}^{n+1} - \sum_{F' \subset \partial K} (M_K^{-1})_{F',F} t p_{K,F'}^{n+1} - \sum_{F' \subset \partial K'} (M_{K'}^{-1})_{F',F} t p_{K',F'}^{n+1} = 0, \quad (2.51)$$

et pour tous les bords ∂K contenant F , on écrit :

$$\sum_{\partial K \supset F} \alpha_{K,F} p_K^{n+1} - \sum_{\partial K \supset F} \sum_{F' \subset \partial K} (M_K^{-1})_{F',F} t p_{K,F'}^{n+1} = 0. \quad (2.52)$$

En remplaçant p_K^{n+1} par son expression obtenue dans l'étape précédente et en regroupant les termes, on obtient :

$$\sum_{\partial K \supset F} \left[\sum_{F' \subset \partial K} \left((M_K^{-1})_{F',F} - \beta_K \frac{\alpha_{K,F'}}{\alpha_K} \right) t p_{K,F'}^{n+1} \right] = \sum_{\partial K \supset F} \left[\alpha_{K,F} (1 - \beta_K) p_K^n + \frac{\beta_K}{\alpha_K} F_K \right], \quad (2.53)$$

et en ajoutant le vecteur qui contient les conditions aux bords, on peut écrire sous forme matricielle :

$$AT_P = b \quad (2.54)$$

où T_P est le vecteur des traces de pressions par face,

$$A = ((A)_{F,F'})_{F,F' \in \partial K} = \sum_{\partial K \supset F} \sum_{F' \subset \partial K} \left((M_K^{-1})_{F',F} - \beta_K \frac{\alpha_{K,F'}}{\alpha_K} \right),$$

$$b = ((b)_F)_{F \in \partial K} = \sum_{\partial K \supset F} \left[\alpha_{K,F} (1 - \beta_K) p_K^n + \frac{\beta_K}{\alpha_K} F_K \right] + g$$

et g est le vecteur qui conditions aux bords.

L'équation (2.54) représente le système global dont les inconnues sont les traces de la pression.

Validation de l'écoulement transitoire implémenté dans 3FLO

a) Cas test : Theis

Ce test consiste à simuler un essai de pompage à débit Q constant sur l'axe de symétrie d'un cylindre et à analyser la variation de charge en différents points du massif, à différentes distances du puits d'essai, respectivement à 1, 5, 10 et 15 m du centre de puits.

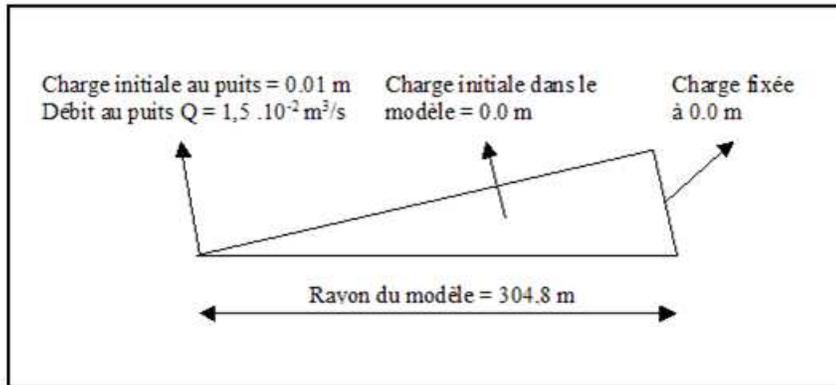


FIG. 2.11: Log de rabattement (m) en fonction du log du temps

Les conditions aux limites et conditions initiales sont illustrées dans la Figure (2.11).

La perméabilité est $10^{-3} \text{ m}^2/\text{s}$, et le coefficient d'emmagasinement est 10^{-3} .

b) Solution analytique de Theis

Le problème de Theis revient à résoudre l'équation de diffusivité en régime transitoire 2D :

$$\Delta h = \frac{S}{\mathcal{K}} \frac{\partial h}{\partial t},$$

où :

- S est le coefficient d'emménagement sans dimension [1],
- \mathcal{K} est la perméabilité [$L^2.T^{-1}$],
- h la charge hydraulique [L],
- t le temps [T].

On suppose que :

- le milieu est infini, radial et symétrique en z , homogène et isotrope,
- la perméabilité est constante (nappe captive, ou nappe libre peu rabattue),
- le problème est plan (surface piézométrique horizontale),
- le forage est vertical en $r = 0$ de petit rayon r_0 , pénètre tout l'aquifère, pompe à débit constant positif Q ,
- l'aquifère est initialement immobile, à $t = 0$ la charge est constante h_0 .

Sous ces hypothèses, la solution ne dépend que de la distance r , et en coordonnées radiales l'équation de la diffusivité s'écrit (cf. [59], [28]) :

$$\begin{cases} \frac{S}{\mathcal{K}} \frac{\partial h}{\partial t} - \frac{\partial^2 h}{\partial r^2} - \frac{1}{r} \frac{\partial h}{\partial r} = 0, & r > r_0, \quad t > 0, \\ h(r, 0) = h_0, & r > r_0, \\ \left(r \frac{\partial h(r, t)}{\partial r} \right)_{r=r_0} = -\frac{Q}{2\pi\mathcal{K}}, & t > 0. \end{cases} \quad (2.55)$$

La solution de ce problème est donné par Theis [86] :

$$h(r, t) = h_0 + \frac{Q}{4\pi\mathcal{K}} \int_{\tau}^{+\infty} \frac{e^{-\xi}}{\xi} d\xi,$$

où Q est le débit de pompage [$L^3.T^{-1}$], et $\tau = \frac{r^2}{4\pi t}$.

c) Résultats numériques

On compare la solution analytique de Theis et la solution numérique donnée par la méthode Mixte-Hybride (Figure 2.13). Les courbes continues (la1, la5, la10, la15) correspondent à la solution analytique et (lh1, lh5, lh10, lh15) à la solution numérique.

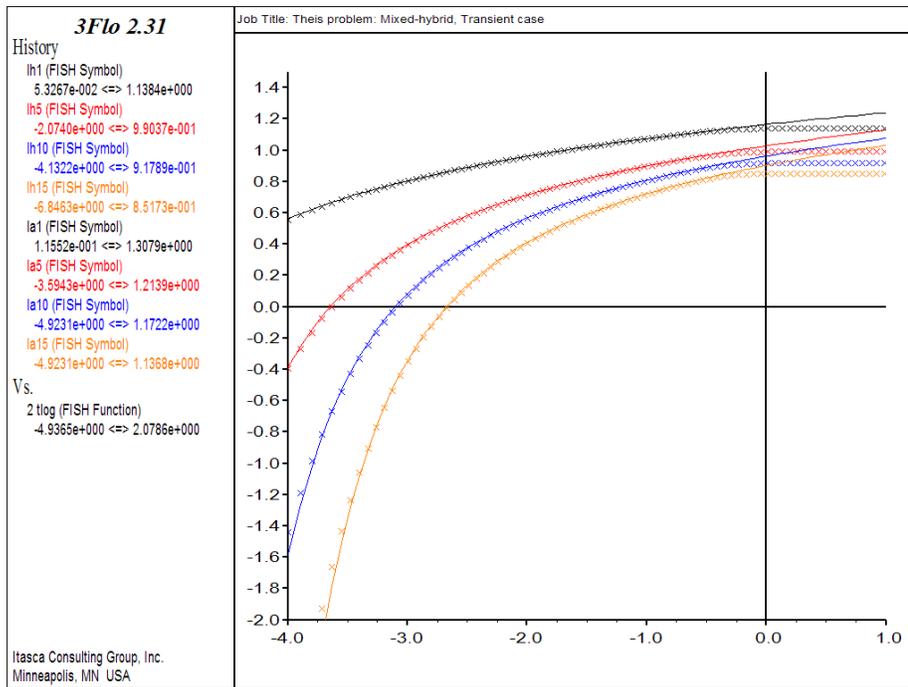


FIG. 2.12: Rabattement (m) en fonction du temps, échelles logarithmiques à différentes distances de puits

On constate une bonne concordance entre les résultats numériques et la solution analytique de Theis jusqu'à environ un jour de pompage. Au delà de cette période le problème traité n'est plus équivalent au problème Theis, qui suppose un milieu infini.

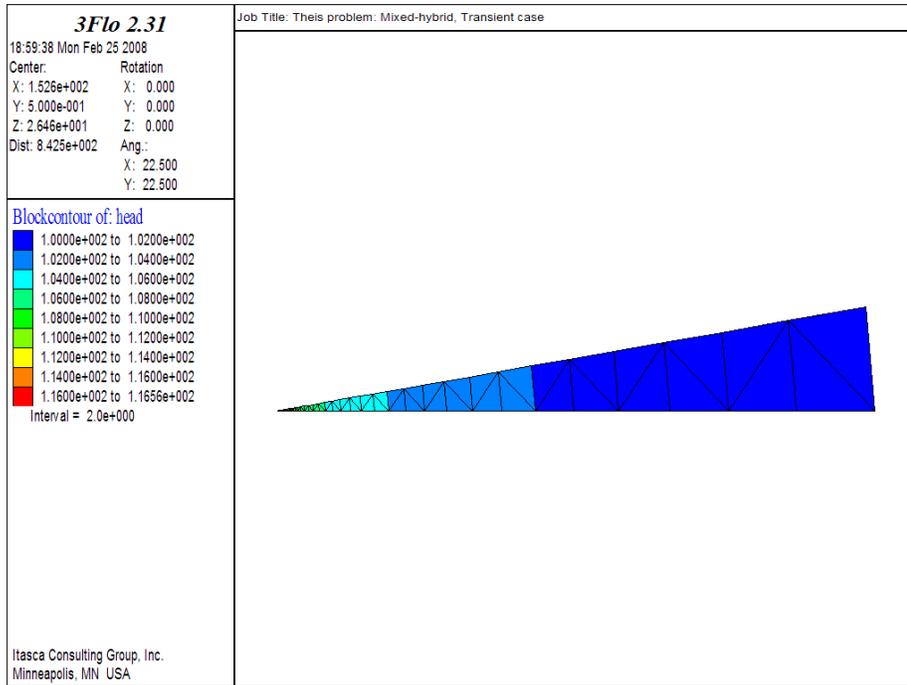


FIG. 2.13: Allure de la charge induite par un pompage à débit constant

6.2 Transport : l'advection par VF et la diffusion par EFMH

Décomposition d'opérateurs

Étape d'advection :

Soit Δt le pas de temps de diffusion, il sera divisé en M pas de temps d'advection Δt_a tels que $\Delta t = M\Delta t_a$ avec $M \geq 1$, le pas de temps d'advection sera contrôlé par la condition CFL :

$$\Delta t_a \leq \min_K \left(\frac{\phi_K |K|}{\sum_{F^- \subset \partial K} \int_{F^-} u \cdot n_{F^-}} \right), \quad (2.56)$$

où F^- est une face du tétraèdre K par laquelle le flux est entrant.

L'intervalle $[t_n, t_{n+1}]$ est divisé en M intervalles $[t_{n,m}, t_{n,m+1}]$, $m = 0, \dots, M-1$, avec $t_{n,0} = t_n$, $t_{n,M} = t_{n+1}$. On note par $c^{n,m}$ l'approximation de c à l'instant $t^{n,m}$ avec $c^{n,0} = c^n$. Le pas de temps d'advection est $\Delta t_a = \Delta t/M$.

Tout d'abord, on résout M fois une équation d'advection explicitement par la méthode de volumes finis décentrée amont.

Pour $m = 0, \dots, M - 1$: $\phi \frac{c^{n,m+1} - c^{n,m}}{\Delta t_a} + \nabla (uc) = 0$,

et en utilisant les notations du paragraphe précédent pour la discrétisation de l'équation d'advection, on obtient :

$$\phi_K |K| \frac{c_K^{n,m+1} - c_K^{n,m}}{\Delta t_a} + \sum_{F \subset \partial K} Q_{K,F}^{n,m} = 0, \quad m = 0, \dots, M - 1,$$

ou encore

$$c_K^{n,m+1} = c_K^{n,m} - \frac{\Delta t}{\phi_K M |K|} \sum_{F \subset \partial K} Q_{K,F}^{n,m}, \quad m = 0, \dots, M - 1. \quad (2.57)$$

Étape de diffusion

Une fois $c_K^{n,M}$ obtenue dans l'étape d'advection après M pas de temps d'advection, on calcule c_K^{n+1} par un schéma implicite, en résolvant l'équation de diffusion (2.58) avec le pas de temps Δt par la méthode des éléments finis mixtes-hybrides (section 6.1).

$$\begin{cases} \phi \frac{c^{n+1} - c^{n,M}}{\Delta t} + \nabla (r^{n+1}) = f, \\ r^{n+1} = -D \nabla c^{n+1} \end{cases} \quad (2.58)$$

En procédant ainsi, on peut résoudre séparément l'advection et la diffusion avec deux échelles différentes de pas du temps. Pour raison de stabilité, on doit choisir un pas de temps d'advection petit qui vérifie la condition CFL, mais on peut se permettre un pas de temps plus grand pour la diffusion, mais ce pas de temps de diffusion ne doit pas être trop grand afin de ne pas trop perdre de précision.

La diffusion par la méthode mixte hybride

Pour discrétiser l'équation de dispersion (2.7) par la méthode EFMH, nous introduisons une nouvelle variable $r = -D \nabla c$ qui correspond au flux dispersif. L'équation (2.7) s'écrit donc de la façon suivante :

$$\begin{cases} \phi \frac{\partial c}{\partial t} + \nabla \cdot r = f, \\ r = -D \nabla c. \end{cases} \quad (2.59)$$

L'approximation par les EFMH du système (2.59), décrivant le transport purement dispersif est identique à celle utilisée pour le système décrivant l'écoulement transitoire.

L'advection par la méthode des volumes finis décentrée en amont

Cette partie décrit un schéma basé sur une méthode de volumes finis pour résoudre le problème purement advectif (étape d'advection dans la décomposition d'opérateurs). On considère l'équation d'advection suivante :

$$\begin{cases} \phi \frac{\partial c(x, t)}{\partial t} + \nabla \cdot (u(x, t)c(x, t)) = 0, \\ c(x, 0) = c_0(x). \end{cases} \quad (2.60)$$

Soit Ω le domaine d'application et \mathcal{T} le maillage de discrétisation. On intègre l'équation (2.60) sur un élément K de \mathcal{T} , on obtient :

$$\int_K \phi \frac{\partial c(x, t)}{\partial t} dx + \int_{\partial K} u(x, t) \cdot \nu_K(x) c(x, t) dx = 0, \quad (2.61)$$

où ν_K est le vecteur normal extérieur de ∂K . Soit Δt le pas de temps tel que $t_n = n\Delta t$ pour $n \in \mathbb{N}$. Avec une discrétisation en temps par un schéma d'Euler explicite, on approche la solution $c^n(x)$ de l'équation (2.60) à t_n par :

$$\frac{1}{\Delta t} \int_K \phi (c^{n+1}(x) - c^n(x)) dx + \int_{\partial K} u(x, t_n) \cdot \nu_K(x) c^n(x) dx = 0, \quad \forall n \in \mathbb{N}, \quad (2.62)$$

avec $u^0(x) = u(x, 0) = u_0(x)$. Soit c_K^n la concentration inconnue, constante dans l'élément K , ϕ_K la porosité moyenne dans K , et soit $\nu_{K,F}$ la normale à la face F du bord ∂K dirigé vers l'extérieur de K . Le deuxième intégrale dans (2.62) peut s'écrire sous forme :

$$\int_{\partial K} u(x, t_n) \cdot \nu_K(x) c^n(x) dx = \sum_{F \in \partial K} \int_F u(x, t_n) \cdot \nu_{K,F} c^n(x) dx. \quad (2.63)$$

Pour $F \subset \partial K$, soit

$$Q_{K,F}^n = \int_F u(x, t_n) \cdot \nu_{K,F} dx.$$

Ce terme représente le flux à travers la face F , et c'est la quantité calculée par les EFMH dans l'écoulement.

Par un décentrage en amont, chaque terme de la somme du second membre de (2.63) est discrétisé par :

$$q_{K,F}^n = \begin{cases} Q_{K,F}^n c_K^n & \text{si } Q_{K,F}^n \geq 0, \\ Q_{K,F}^n c_{K'}^n & \text{si } Q_{K,F}^n < 0, \end{cases} \quad (2.64)$$

où K' est l'élément voisin de K de face commune F . Avec cette approximation, le problème (2.60) peut s'écrire sous forme :

$$\begin{cases} \phi_K \frac{|K|}{\Delta t} (c_K^{n+1} - c_K^n) + \sum_{F \in \partial K} q_{K,F}^n = 0, & \forall K \in \mathcal{T}, \\ c_K^0 = \int_K c_0(x) dx, \end{cases} \quad (2.65)$$

où $|K|$ est la mesure de l'élément K . Ce schéma est localement conservatif, pour chaque face F entre deux éléments K et K' , on a $Q_{K,F}^n = -Q_{K',F}^n$. Cette propriété est importante pour plusieurs champs d'applications.

6.3 Mise en oeuvre dans 3FLO pour la partie diffusive du transport

Le logiciel 3FLO est un code de calcul d'écoulement et de transport 3D. Il est adapté aussi bien pour les études hydrogéologiques dans les milieux fracturés que pour les études géochimiques dans les milieux poreux.

Pour mettre en oeuvre la partie diffusive du transport, on a d'abord modifié le code 3Flo pour pouvoir simuler un écoulement transitoire avec la méthode mixte hybride, et les résultats numériques ont été validés par une solution analytique sur le cas test Theis, cf. la section 6.1 dans le paragraphe validation.

Comme la résolution de la partie diffusive du transport dans la décomposition d'opérateurs est similaire à un problème d'écoulement transitoire, alors sa résolution revient à adapter le développement précédent pour simuler un problème de diffusion pure. Le but était d'aboutir à un code unique pour traiter les deux situations.

La difficulté pour modifier le code est venue de ce que les variables faisaient explicitement référence aux quantités physiques (charge, pression, perméabilité, ..), les méthodes de résolution étaient génériques et n'ont pas été modifiées. Il fallait donc trouver un mécanisme qui permette de faire agir ces méthodes génériques sur des données qui peuvent être de type différents, et éviter de dupliquer le code.

Le principe

Pour pouvoir distinguer entre l'écoulement et la diffusion, on a introduit une classe abstraite `MainResolution` et deux classes dérivées `Flow` et `Diff` qui héritent de la classe `MainResolution`. La classe `Flow` est spécifique pour une résolution d'écoulement stationnaire ou transitoire et la classe `Diff` pour un problème de diffusion avec la méthode Mixte hybride ou Galerkin. La classe `MainResolution` contient les méthodes de résolution, et délègue aux sous-classes l'accès aux données, via le mécanisme des fonctions virtuelles (voir diagramme (2.14)).

Restructuration du code

Diagramme des classes

La classe `MainResolution` contient la définition des fonctions Mixte hybride et Galerkin (`GResolution()`, `Gassembly()`, `MHResolution()`, `MHassem()`). Dans ces

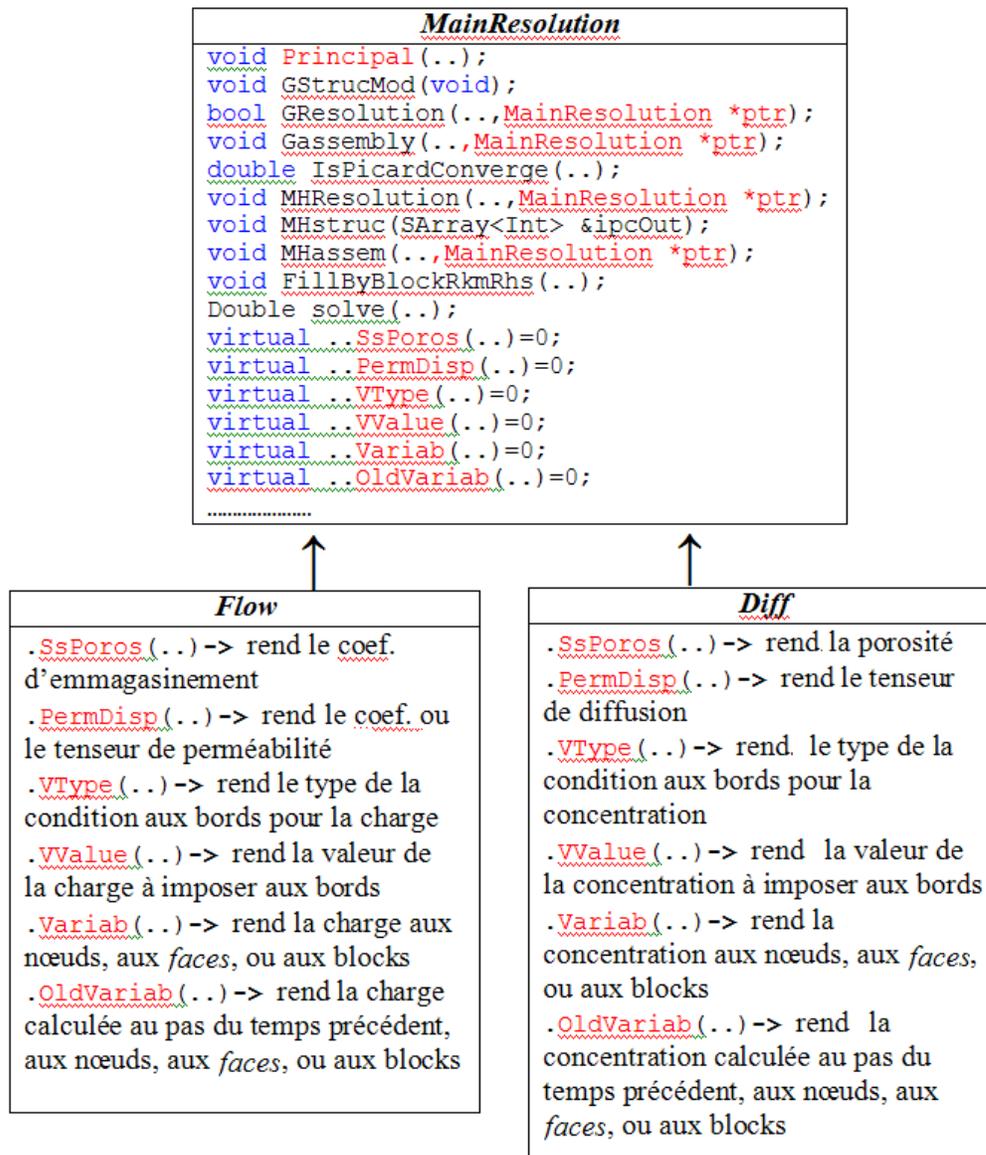


FIG. 2.14: Diagramme des classes

fonctions, on a remplacé toutes les variables spécifiques pour une résolution d'écoulement par les fonctions virtuelles (mentionnées dans le diagramme (2.14)) . La fonction *Variab()* ou *OldVariab()* renvoie la charge pour un calcul de l'écoulement et la concentration pour un calcul de diffusion. Dans une fonction *Principal()* de la classe mère, on fait une mise à jour de *OldVariab* à chaque pas du temps et on résout le problème avec Galerkin ou Mixte hybride.

La classe `HeadTank` est remplacée par la classe `GeneralTank` pour contenir en plus des informations sur la charge (les conditions aux bords sur la charge, *head*, *oldhead*), les données et les informations pour la concentration.

Avec cette réorganisation du code, il devient plus simple et propre de distinguer entre une résolution d'écoulement ou de diffusion et il n'est pas nécessaire de créer et d'utiliser partout un indicateur qui distingue les deux ; la distinction se fait maintenant par un pointeur de la classe `MainResolution`,

```
MainResolution *ptrF=new Flow();
MainResolution *ptrD=new Diff();
```

Dans la boucle principale du temps, on appelle la fonction *Principal()* une fois avec l'objet *ptrF* (*ptrF* → *Principal()*) pour une résolution d'écoulement et une autre fois avec l'objet *ptrD* (*ptrD* → *Principal()*) dans la fonction `TransportSplitting()` pour une résolution du problème de diffusion.

Autres modifications dans le code

- Mise en oeuvre de la partie advective par la méthode des volumes finis.
- Schéma de temps splitting entre advection et diffusion.
- Calcul de la vitesse de Darcy et du tenseur de diffusion.
- Adaptation de la méthode de Galerkin pour la réutiliser pour un calcul de diffusion.
- Distinction entre les concentrations calculées à partir de la quantité de particules et les concentrations dues à un gradient de concentration (accès au fichier de données "fish", graphiques, commandes d'initialisation).
- Ajout de la capacité de stocker la charge et la concentration dans le block et deux commandes *bcConc* et *bcBconc* pour tracer le contour des concentrations ou `blockContour concentrations`.
- Ajout des fonctions fish *block_conc*, *pipe_conc*, *node_conc*, *bl_head*.
- Initialisation de la concentration et prise en compte des conditions aux bords sur la concentration due à un gradient.
- Ajout de la commande "*set transport*" pour distinguer entre un calcul d'écoulement ou du transport.

7 Validation numérique du transport avec la solution analytique

7.1 Cas test académique

Pour ce cas test académique, on résout un problème d'advection-diffusion [80], [38], et on veut comparer les résultats de concentrations donnés par le code splitting implémenté dans 3Flo, et la solution analytique bidimensionnelle [38] qui correspond aux conditions suivantes

- le domaine est homogène ;
- l'écoulement est monodimensionnel, avec une vitesse parallèle à l'axe (ox) ($v=(1,0,0)$) ;
- La concentration initiale est nulle ;
- La condition au bord est une condition de Dirichlet, une concentration imposée sur une partie du bord à gauche du domaine, $c(0,y , t>0)=c_0$ ($c_0=10$) si $|y|<a$ ($a=8$).

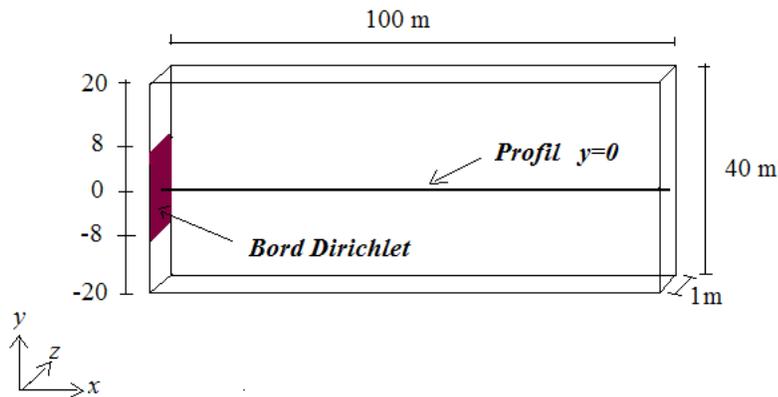


FIG. 2.15: Géométrie

Pour les autres données [80] du problème, on prend un domaine de longueur 100 et de profondeur 40 (et de largeur 1 pour la simulation 3D dans 3Flo). Le maillage utilisé dans 3Flo est un maillage tétraédrique (avec $\Delta x = 2$, $\Delta y = 2$, $\Delta z = 1$) Deux régimes testés, régime diffusif avec $Pe = 1$ (pour une diffusion longitudinale $D_L = 2$ et transversale $D_T = 0.2$) , et régime advectif avec $Pe = 10$ (pour les données : $D_L = 0.2$ et $D_T = 0.02$).

7.2 La solution analytique

La solution analytique avec ces conditions est donnée selon Leij Feike et Dane [38] par :

$$c(x, y, t) = \frac{xc_0}{(16\pi D_L)^2} \int_0^t \tau^{-\frac{3}{2}} \left\{ \operatorname{erf} \left[\frac{a+y}{(4D_T\tau)^{\frac{1}{2}}} \right] + \operatorname{erf} \left[\frac{a-y}{(4D_T\tau)^{\frac{1}{2}}} \right] \right\} \exp \left[- \left(\frac{x-v\tau}{(4D_L\tau)^{\frac{1}{2}}} \right)^2 \right] d\tau.$$

Pour calculer l'intégrale, on va utiliser la formule de Gauss-Chebyshev basée sur les polynômes de Chebyshev d'ordre 1,

$$\int_a^b f(x)dx \simeq \frac{b-a}{2} \sum_{k=1}^n w(\xi_k) \sqrt{1-\xi_k^2} f \left(\frac{b-a}{2} \xi_k + \frac{b+a}{2} \right),$$

où

$$\xi_k = \cos \frac{(2k-1)\pi}{2n} \quad \text{et} \quad w(\xi_k) = \frac{\pi}{n} \quad \text{pour} \quad k = 1, \dots, n,$$

et

$$f(\tau) = \tau^{-\frac{3}{2}} \left\{ \operatorname{erf} \left[\frac{a+y}{(4D_T\tau)^{\frac{1}{2}}} \right] + \operatorname{erf} \left[\frac{a-y}{(4D_T\tau)^{\frac{1}{2}}} \right] \right\} \exp \left[- \left(\frac{x-v\tau}{(4D_L\tau)^{\frac{1}{2}}} \right)^2 \right].$$

Avec un changement de variables, on obtient :

$$\int_0^t f(x)dx \simeq \frac{t}{2} \sum_{k=1}^n \frac{\pi}{n} \sqrt{1-\xi_k^2} f \left(\frac{1+\xi_k}{2} t \right),$$

$$\text{comme } \frac{1+\xi_k}{2} = \frac{1}{2} \left(1 + \cos \frac{(2k-1)\pi}{2n} \right) = \cos^2 \frac{(2k-1)\pi}{4n}$$

alors

$$\int_0^t f(x)dx \simeq \frac{t}{2} \sum_{k=1}^n \frac{\pi}{n} \sqrt{1-\xi_k^2} f \left(t \cos^2 \frac{(2k-1)\pi}{4n} \right),$$

soit

$$\int_0^t f(x)dx \simeq \frac{t}{2} \sum_{k=1}^n \frac{\pi}{n} \sin \frac{(2k-1)\pi}{2n} f \left(t \cos^2 \frac{(2k-1)\pi}{4n} \right).$$

Avec cette approximation de l'intégrale, la solution peut s'écrire sous la forme

$$c(x, y, t) \simeq \frac{xtc_0}{(32\pi D_L)^2} \sum_{k=1}^n \frac{\pi}{n} \sin \frac{(2k-1)\pi}{2n} f \left(t \cos^2 \frac{(2k-1)\pi}{4n} \right).$$

7.3 Comparaison des résultats

Dans cette partie, on compare les résultats numériques et analytiques du transport pour les deux régimes advectif et diffusif, et on discute l'influence du pas de temps d'advection ainsi du pas de temps de diffusion sur la solution numérique.

Advection dominante

Les courbes d'iso-valeurs (Fig. 2.16) représentent le panache des concentrations à $t=50$, à droite la solution analytique et à gauche la solution numérique donnée par le code `splitting (3Flo)`, le pas de temps de diffusion est $\Delta t = 10$ et le nombre de pas d'advection pour un pas de diffusion vaut $M = 18$.

Comme on s'y attend, le panache se déplace linéairement suivant la vitesse et le sens d'écoulement, et les résultats numériques sont très proches de la solution analytique.

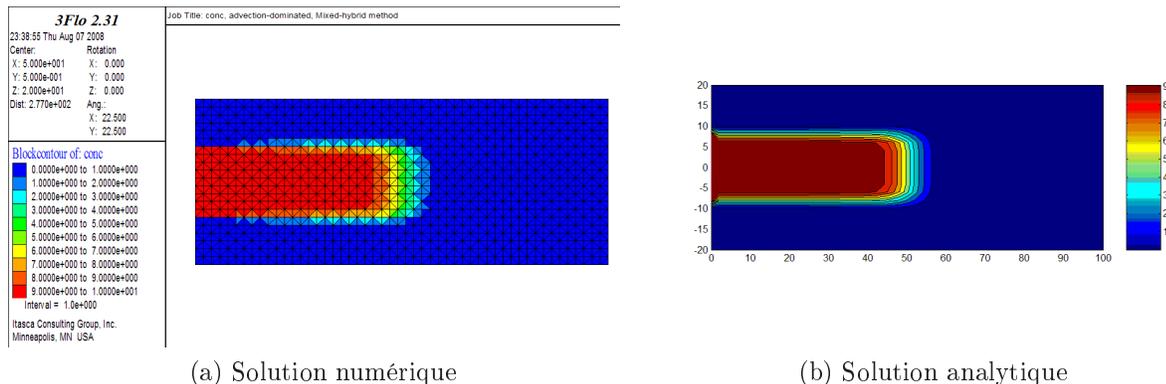


FIG. 2.16: Panache de concentration à $t=50$

Pour comparer plus précisément ces résultats, on va regarder le profil de la concentration suivant l'axe l'écoulement (ox). On trace sur la même figure (Fig. 2.17) l'évolution de la concentration numérique et de la concentration analytique pour le profil $y=0$, à cinq instants différents $t=10, 20, 30, 40$, et 50 . On constate que la concordance est excellente.

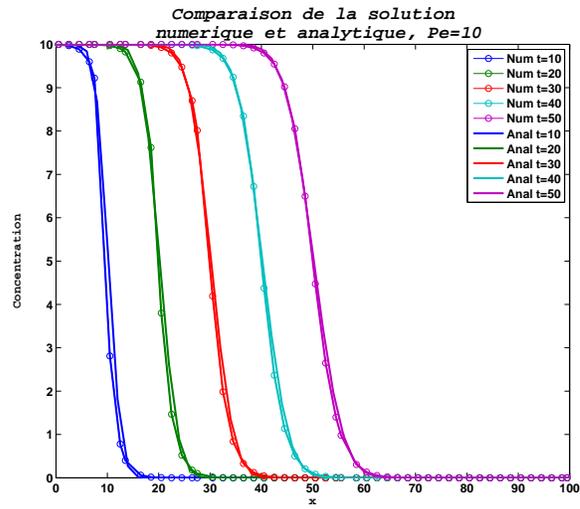


FIG. 2.17: Comparaison de la concentration numérique et de la solution analytique pour le profil $y=0$, et $Pe=10$.

L'effet du pas de temps d'advection sur la solution

Pour voir l'effet de la condition CFL et le choix du pas de temps d'advection, on fixe pour ce cas test, le pas de temps de diffusion $\Delta t = 10$, le nombre de Peclet $Pe = 10$, le pas de temps $d_{tmax} = 0.65359$. On fait varier la valeur de CFL, ce qui veut dire varier le pas du temps d'advection (Tableau à droite), et on compare les solutions numériques obtenues avec la solution analytique (Fig. 2.18).

CFL	d _t c	M
0.1	0.06493	154
1	0.625	16
1.1	0.71428	14

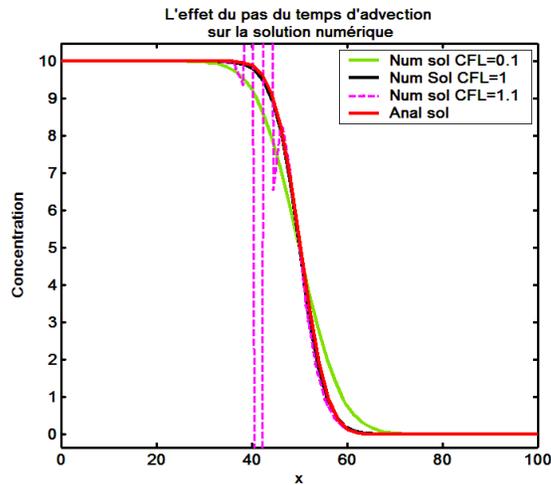


FIG. 2.18: Concentration pour différentes valeurs de CFL

Le respect de la condition CFL est un critère nécessaire pour la stabilité d'un schéma numérique, on constate la présence d'oscillations numériques lorsqu'on dépasse sa valeur limite 1. Avec la valeur CFL qui vaut 1, le résultat numérique coïncide avec la solution analytique et on remarque une diffusion numérique qui apparaît dès que la valeur CFL s'éloigne de 1.

Diffusion dominante

Ce cas test représente le modèle du transport pour le régime diffusif ($Pe = 1$), les courbes d'iso-valeurs (Fig. 2.19) représentent le panache des concentrations à $t=50$, à droite la solution analytique et à gauche la solution numérique.

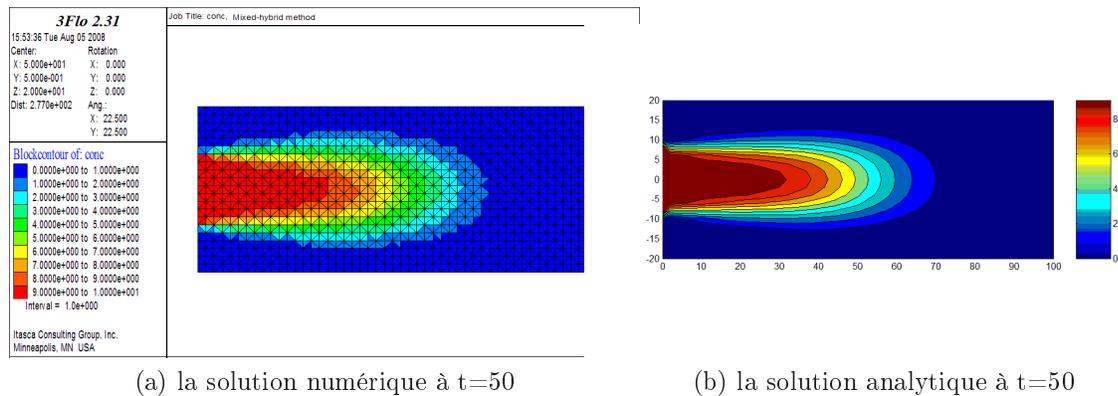


FIG. 2.19: Panache de concentration

La figure (Fig. 2.20) représente l'évolution de la concentration numérique et analytique pour le profil $y=0$, à cinq instants différents $t=10, 20, 30, 40$, et 50 . Les résultats numériques tracés sur la figure (Fig. 2.20a) sont donnés pour un grand pas de temps de diffusion $\Delta t = 10$ pour 18 pas d'advection, et les résultats numériques de la figure (Fig. 2.20b) sont données pour un petit pas de temps de diffusion $\Delta t = 1.5$ pour 3 pas d'advection, et ces résultats prouvent une très bonne conformité avec la solution analytique.

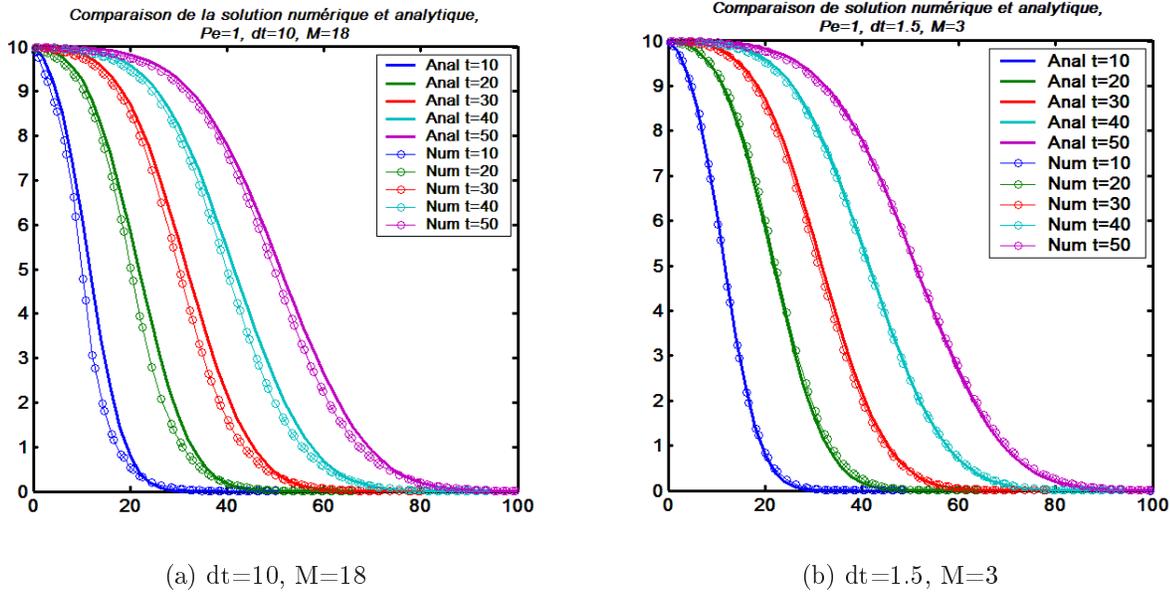


FIG. 2.20: Comparaison de la concentration numérique et analytique pour le profil $y=0$, et $Pe=1$

Pour le régime diffusif, l'erreur d'approximation diminue avec le pas du temps de diffusion.

8 Conclusion

La résolution de l'équation d'advection-diffusion en milieux poreux nécessite la connaissance préalable du champ de vitesse. L'utilisation de la méthode EFMH pour le calcul du champ d'écoulement assure une description de l'écoulement beaucoup plus précise que les méthodes d'EF de Lagrange.

La résolution numérique de l'équation du transport est un problème relativement difficile, notamment quand le mécanisme de l'advection est prédominant ou lorsque la simulation doit être effectuée sur une très longue période de temps. Sa résolution par les méthodes classiques n'est satisfaisante que pour une discrétisation très fine du domaine d'espace et du temps.

La technique de séparation d'opérateurs est utilisée pour discrétiser l'équation d'advection-diffusion. La partie diffusive est approchée en utilisant la méthode EFMH et la partie advective est approchée par la méthode VF. L'intérêt de cette technique est de traiter des processus différents comme l'advection, la diffusion, la chimie, etc. avec des méthodes numériques appropriées, ainsi d'économiser du temps de calcul en prenant des échelles de temps différents pour la diffusion et l'advection.

Chapitre 3

Résolution du modèle d'équilibre chimique

Dans ce chapitre, nous rappellerons les éléments du traitement numérique des réactions chimiques. Les ingrédients essentiels seront la loi d'action de masse, et la loi de conservation de la masse.

Dans le cadre du transport réactif, le rôle du module chimique sera de calculer quelle proportion de chaque espèce se trouve en solution, et quelle partie se trouve sous forme solide. Il est donc important de distinguer les réactions homogènes et hétérogènes. Nous donnerons quelques éléments sur la modélisation des réactions de sorption, en indiquant comment on peut les ramener, du point de vue du calcul, à des lois d'action de masses.

Dans cette thèse nous nous restreignons aux réactions en équilibre, ce que nous justifierons après une rapide étude des réactions cinétiques. Nous introduirons les notions de composants et d'espèces secondaires, dues à Morel [65], et qui permettent de formuler de manière concise un système chimique à l'équilibre.

Cette formulation conduit à un système d'équations non-linéaires. Après avoir discuté de l'existence et de l'unicité des solutions, nous étudierons la résolution numérique de ce système par la méthode de Newton. Nous validerons cette étude sur des cas tests.

Nous indiquerons la modélisation des réactions de sorption, qui sont indispensables pour l'étude du modèle couplé au chapitre suivant, et les modifications à apporter à la méthode vue précédemment pour les prendre en compte.

1 Modélisation des réactions chimiques

Le but d'un module de spéciation chimique est de calculer les concentrations des différentes espèces présentes dans le système considéré, à partir de la connaissance des réactions, et des éléments chimiques en cause.

On peut classer les réactions chimiques selon deux critères (voir [72]) : en

fonction du caractère homogène ou non de la réaction, ou en fonction de la vitesse de la réaction.

1.1 Réactions homogènes ou hétérogènes

Une première classification des réactions est entre réactions homogènes (dans une même phase), et réactions hétérogènes (faisant intervenir deux ou plusieurs phases). Sans donner de définition précise, notons qu'une phase est un ensemble homogène dans sa composition.

Dans le cadre des modèles de milieux souterrains, les réactions homogènes sont les réactions en phase *aqueuse*, c'est-à-dire que d'une part l'eau intervient de manière essentielle, et que d'autre part la matrice rocheuse et les espèces minérales n'interviennent pas. On trouvera dans ce cadre, outre la dissociation de l'eau, la complexation, les réactions acido-basiques, les réactions d'oxydo-réduction, ...

Les réactions hétérogènes (dans le contexte géochimique) sont principalement les réactions de précipitation-dissolution, et les réactions de sorption. Rubin [72] distingue entre ces deux types de réaction en considérant que les premières sont de type "classique", alors que les secondes se produisent à l'interface du milieu aqueux avec la roche. Les réactions de précipitation-dissolution impliquent l'eau et une (ou plusieurs) espèce minérale, qui peut exister sous forme solide, ou sous forme dissoute. La réaction ne se produit que si un seuil, nommé *produit de solubilité* est atteint. Une conséquence importante du point de vue de la modélisation est que ces réactions conduisent à des modèles non-différentiables. Nous ne les prendrons pas en compte dans cette thèse.

Les réactions de sorption sont le principal mécanisme par lequel les espèces chimiques dissoutes interagissent avec la matrice rocheuse. Il existe deux modèles principaux pour décrire ces échanges (nous y reviendrons brièvement au paragraphe 4) : la complexation de surface, et l'échange d'ions. Dans le cadre de ce travail, au delà des différences entre ces deux approches, nous considérerons qu'elles conduisent toutes deux à des réactions chimiques qui rentrent dans le formalisme général des lois d'action de masse.

1.2 Réactions cinétiques ou à l'équilibre

L'autre axe de classification des réactions chimiques est celui de leur vitesse, ou de leur caractère réversible ou irréversible. Dans le cas de la géochimie, il est naturel de comparer cette vitesse aux vitesses caractéristiques du transport. Quand la vitesse de réaction est lente par rapport à celle du transport, ou du même ordre que celle-ci, on doit prendre en compte la cinétique de la réaction, qui est alors irréversible. Dans le cas contraire, le système est réversible, on peut considérer qu'il est à *l'équilibre instantané à chaque instant*, et qu'une modélisation par des réactions à l'équilibre est adéquate.

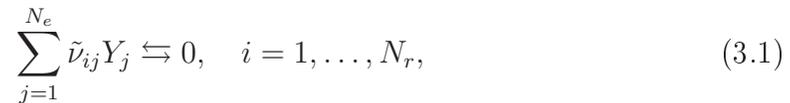
L'importance des réactions cinétiques dépend du type d'application envisagée. Les réactions en phase aqueuse peuvent être considérées à l'équilibre, et il en est de même des réactions de sorption. Les applications incluant des réactions biochimiques doivent prendre en compte leur caractère cinétique [43], et il est souvent nécessaire de prendre en compte la cinétique des réactions minérales (une des raisons pour ignorer ce point est d'ailleurs la difficulté de disposer de données expérimentales sur les vitesses de réaction).

Contrairement au modèle d'équilibre, régi par la thermodynamique, et que nous décrirons au paragraphe 2, il existe une grande variété de modèles cinétiques. Citons par exemple les modèles de biodégradation [43], les modèles prenant en compte le produit de solubilité des minéraux [81], et les modèles basés sur la loi d'action de masse. À la fois en raison de cette difficulté, et conformément au choix de ne pas prendre en compte les minéraux, dans cette thèse nous nous restreignons aux réactions à l'équilibre.

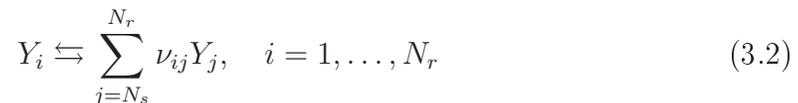
2 Modélisation des réactions d'équilibre

2.1 Espèces composantes et secondaires

Nous considérons donc maintenant un système chimique composé uniquement de réactions à l'équilibre. À ce stade, nous ne distinguerons pas entre les réactions homogènes ou hétérogènes, cela fera l'objet de la section 4. Nous considérons donc un système composé de N_e espèces, réagissant entre elles par N_r réactions :



où $\tilde{\nu}$ est la matrice des *coefficients stoechiométriques*. Il est naturel de supposer, d'une part que $N_r \leq N_e$, et d'autre part que la matrice $\tilde{\nu}$ est de rang maximal, soit N_r . En effet, dans le cas contraire, des réactions seraient redondantes. Notons que cette hypothèse peut tomber en défaut dans le cas des modèles cinétiques, pour lesquels il peut exister des chemins réactionnels distincts entre deux états. Mettons la matrice $\tilde{\nu}$ sous forme échelon [63] (cette réduction n'est pas unique). Comme la matrice est supposée de rang maximum, cette forme échelon est $[I \quad -\nu]$ où la matrice identité est de taille $N_r \times N_r$, et ν est $N_s - N_r \times N_r$. Le système chimique s'écrit alors (après une éventuelle renumérotation) sous la forme

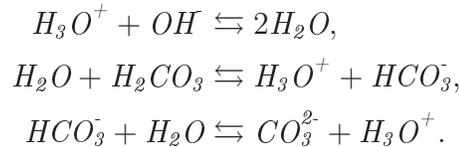


Il est commode d'isoler les $N_s - N_r$ dernières composantes du système, et de partitionner $Y = \begin{bmatrix} x \\ c \end{bmatrix}$, avec x de taille N_r et c de taille $N_s - N_r$. Les espèces présentes

dans c sont appelées les espèces *composantes* du système, et celles dans x sont dites espèces secondaires. On notera $N_c = N_s - N_r$ le nombre de composantes. Avec ces notations, le système chimique devient finalement

$$x_i \rightleftharpoons \sum_{j=1}^{N_c} \nu_{ij} c_j, \quad i = 1, \dots, N_r. \quad (3.3)$$

Exemple 1. *Considérons le système chimique de la dissociation de l'acide carbonique, formé par les réactions suivantes :*



Les espèces en présence sont OH , HCO_3^- , CO_3^{2-} , H_2O , H_3O^+ , et H_2CO_3 , on a donc $N_s = 6$, $N_r = 3$, et la matrice stoechiométrique est

$$\tilde{\nu} = \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1 & -1 & 0 \end{bmatrix}$$

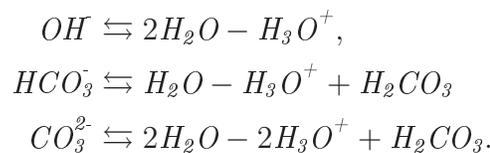
Comme la matrice est de rang 3, il y a plusieurs choix possibles pour les espèces composantes. Nous choisissons (arbitrairement) H_2O , H_3O^+ , et H_2CO_3 . Une forme échelon pour $\tilde{\nu}$ est

$$\begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & -2 & 2 & -1 \end{bmatrix}.$$

On voit donc que

$$\nu = \begin{bmatrix} 2 & -1 & 0 \\ 1 & -1 & 1 \\ 2 & -2 & 1 \end{bmatrix}.$$

et l'on peut exprimer les espèces secondaires, OH , HCO_3^- et CO_3^{2-} :



Ainsi, les espèces composantes sont celles à partir desquelles on peut exprimer, de façon unique, les autres espèces présentes dans le système. Il s'agit en fait de trouver une base de l'espace vectoriel engendré par les colonnes de la matrice stoechiométrique $\tilde{\nu}$.

2.2 Mise en équation

Le but d'un calcul d'équilibre est de trouver la répartition des différentes espèces chimiques. Les lois régissant cet équilibre sont celles de la thermodynamique classique, en particulier l'équilibre est obtenu lorsque la composition du système réalise le minimum de l'énergie libre

$$G = \sum_i n_i \mu_i$$

où n_i est le nombre de mûles de l'espèce i et μ_i est son potentiel chimique. Le potentiel chimique de l'espèce i est donné en fonction de son *activité* a_i par

$$\mu_i = \mu_i^0 + RT \ln \frac{a_i}{a_i^0},$$

où μ_i^0 et a_i^0 sont, respectivement, le potentiel chimique et l'activité de l'espèce considérée, dans des conditions de référence. Pour toutes ces notions, voir par exemple les livres [81] ou [82]. Dans ce dernier ouvrage, on montre que la minimisation de l'énergie libre est équivalente à la formulation classique par loi d'action de masse et conservation de la masse, et c'est cette formulation que nous allons décrire maintenant.

Pour chaque réaction chimique, l'équilibre s'exprime par la loi d'action de masse : le produit (avec des puissances algébriques) des activités des espèces participant à la réaction est égal à une constante, appelée *constante d'équilibre* de la réaction. Cette constante (qui peut, en principe, se calculer à partir du potentiel chimique de référence) est une donnée pour chaque réaction chimique (elle peut varier en fonction de la température). On trouve ces constantes dans les bases de données de thermodynamique. Pour chaque réaction telle que (3.3), on écrira donc la relation algébrique :

$$\{x_i\} = K_i \prod_{j=1}^{N_c} \{c_j\}^{\nu_{ij}}, \quad i = 1, \dots, N_s \quad (3.4)$$

en notant $\{c_j\}$ (resp. $\{x_i\}$) l'activité du composant j (resp. de l'espèce secondaire i).

Une autre série d'équations est fournie par la loi de conservation des espèces. Si l'on considère un système fermé, la quantité de chaque espèce présente est invariante, et se retrouve donc soit sous la forme du composant considéré, soit sous la forme de l'une des espèces secondaires. On notant T_i le total du composant i (supposé connu pour un système fermé), on a donc

$$T_j = [c_j] + \sum_{i=1}^{N_s} \nu_{ij} [x_i], \quad j = 1, \dots, N_c \quad (3.5)$$

en notant cette fois $[c_j]$ (resp. $[x_i]$) la concentration du composant j (resp. de l'espèce secondaire i).

Il reste à préciser la relation entre activité et concentration. Il s'agit là d'un problème de modélisation relativement délicat, et pour lequel il n'existe pas de consensus dans la communauté des chimistes. Pour une solution idéale, c'est-à-dire dans la limite d'une dilution infinie (en pratique, lorsque les concentrations sont faibles devant 1), on peut confondre activité et concentration. On parle donc de correction d'activité.

Pour prendre en compte cette correction, on introduit une fonction $\gamma(y)$ (qui est donc une fonction de l'ensemble des concentrations, nous notons $[y] = \begin{bmatrix} [x] \\ [c] \end{bmatrix}$), appelée *coefficient d'activité*, qui est telle que $[a_k] = \gamma(y)[y_k]$, et qui vérifie donc $\lim_{y \rightarrow 0} \gamma(y) = 1$. La plupart des modèles utilisés, comme le modèle de Debye-Hückel, ou le modèle de Davies (voir [81]) font intervenir la force ionique définie comme

$$I = 1/2 \sum_k [y_k] Z_k^2$$

où $[y_k]$ est la concentration de l'espèce (composante ou secondaire) k , et Z_k est sa charge. On voit que cette définition est globale, et que les modèles d'activité vont donc coupler entre elles les différentes concentrations.

Dans cette thèse, nous ne considérerons que la cas des solutions idéales, et ne prendrons donc pas en compte de modèle de correction d'activité. Cela revient à confondre activité et concentration, et pour simplifier les notations, nous noterons donc par c_j ou x_i la concentration de l'espèce correspondante. Les équations (3.4) et (3.5) se réécrivent donc sous la forme

$$x_i = K_i \prod_{j=1}^{N_c} c_j^{\nu_{ij}}, \quad i = 1, \dots, N_s \quad (3.6)$$

$$T_j = c_j + \sum_{i=1}^{N_s} \nu_{ij} x_i, \quad j = 1, \dots, N_c \quad (3.7)$$

Ces équations forment un système d'équations non-linéaires pour les concentrations des différentes espèces constituant le système chimique étudié. Avant de passer à la résolution numérique de ce système, nous allons introduire une manière synthétique de résumer ce système, qui est courante dans la littérature géochimique, et que nous utiliserons dans la suite de ce travail. Nous ferons ensuite quelques commentaires concernant l'existence et l'unicité des solutions du système (3.6).

2.3 Le tableau de Morel

Une représentation synthétique du système chimique est fournie par le tableau de Morel [65]. On écrit simplement la matrice ν (parfois sous la forme $\begin{bmatrix} I \\ \nu \end{bmatrix}$ pour faire figurer toutes les espèces), en indiquant au dessus de chaque colonne (resp. à gauche

de chaque ligne) le nom de la l'espèce composante (resp. secondaire) correspondante, en bas de chaque colonne le total de l'espèce considérée, et à droite de chaque ligne la valeur de la constante d'équilibre.

On lit donc sur le tableau de Morel une réaction chimique (ou de manière équivalente, une loi d'action de masse) par ligne, et une loi de conservation par colonne. Cette présentation est équivalente à la formulation (3.6) donnée plus haut, mais permet, sur un exemple concret de présenter en une seule fois l'ensemble des informations.

De manière générale, le tableau de Morel pour un système chimique s'écrit sous la forme

	c_1	\cdots	c_j	\cdots	c_{N_c}	
x_1	ν_{11}	\cdots	ν_{1j}	\cdots	ν_{1N_c}	K_1
\vdots	\vdots		\vdots		\vdots	\vdots
x_i	ν_{i1}	\cdots	ν_{ij}	\cdots	ν_{iN_c}	K_i
\vdots	\vdots		\vdots		\vdots	\vdots
x_{N_s}	ν_{1N_s}	\cdots	$\nu_{N_s j}$	\cdots	$\nu_{N_s N_c}$	K_{N_s}
Conc. tot.	T_1	\cdots	T_i	\cdots	T_{N_c}	

TAB. 3.1: Tableau de Morel pour un système chimique général

Exemple 2. Nous reprenons les données de l'exemple 1. Le tableau de Morel s'écrit dans ce cas :

	H ₂ O	H ₃ O ⁺	H ₂ CO ³	
OH ⁻	2	-1	0	K_1
HCO ₃ ⁻	1	-1	1	K_2
CO ₃ ²⁻	2	-2	1	K_3
Total	T_1	T_2	T_3	

TAB. 3.2: Tableau de Morel pour la dissociation du carbone

2.4 Existence et unicité pour le système d'équilibre

Avant de discuter des méthodes numériques pour résoudre le système non-linéaire (3.6), nous allons brièvement donner les résultats théoriques existants dans la littérature.

La majorité des travaux discutant de l'existence, et de l'unicité, de solutions pour (3.6) partent de la formulation sous forme de minimisation de l'énergie libre de

Gibbs. Un des premiers articles dans cette veine est celui de Shapiro et Shapley [79]. L'existence d'au moins un minimum est établie sous des hypothèses assez générales (système à plusieurs phases). Il est évidemment nécessaire que les lois de conservation (3.5), qui sont des contraintes du problème de minimisation soient vérifiées, ce qui constitue une condition sur le total T . L'existence découle de la convexité de la fonctionnelle d'énergie. De plus, une condition est donnée pour l'unicité de la solution, qui est en particulier vérifiée dans le cas d'un système à une seule phase, ce qui recouvre le cas traité ici.

Ces considérations sur la convexité sont reprises par de nombreux auteurs (Bethke [9], ainsi qu'un chapitre du livre [82]).

Une autre approche pour l'existence et l'unicité se trouve dans l'article de G. Gnacadja [42] : la formulation (3.6) constitue un système d'équations polynomiales, auxquels l'auteur applique un théorème de point fixe, pour une métrique adaptée au problème.

3 Résolution numérique du système d'équilibre

Le système (3.6) est un système d'équations non-linéaires, dont il n'est naturellement pas possible (en général) de calculer exactement la solution. On doit donc faire appel à une méthode numérique.

Une première difficulté vient de ce que les inconnues naturelles sont les concentrations des différentes espèces. Ces concentrations sont susceptibles de varier sur plusieurs ordres de grandeur, et doivent rester positives pour garder leur signification physique. Ces deux contraintes rendent la résolution numérique délicate. Heureusement, un simple changement de variables permet d'éliminer ces deux difficultés, et a été adopté par la majorité des codes de calcul : on prend comme inconnues les logarithmes des concentrations. Ainsi, les concentrations seront automatiquement positives, et les inconnues du système non-linéaire garderont un ordre de grandeur raisonnable.

Notons $lc = \ln c$ et $lx = \ln x$ ces nouvelles inconnues (on note $\ln y$ le vecteur dont les composantes sont les logarithmes de composantes du vecteur y). Le système (3.6) devient

$$\begin{aligned} lx_i &= \ln K_i + \sum_{j=1}^{N_c} \nu_{ij} lc_j, & i &= 1, \dots, N_s \\ T_j &= c_j + \sum_{i=1}^{N_s} \nu_{ij} x_i, & j &= 1, \dots, N_c, \end{aligned}$$

soit en notation matricielle :

$$\begin{aligned} lx &= lK + \nu lc, \\ T &= c + \nu^T x, \end{aligned} \tag{3.8}$$

avec les relations $lx = \ln x$, $lc = \ln c$. Outre la simplicité des notations, ce changement de variables fait apparaître clairement la nature linéaire des lois d'action de masse, et met en évidence la difficulté du problème à résoudre.

On peut, et c'est le choix que nous avons fait, éliminer les inconnues correspondant aux concentrations des espèces secondaires x . On obtient un système d'équations portant uniquement sur lc :

$$\exp(lc) + \nu^T \exp\{lK + \nu lc\} = T. \quad (3.9)$$

L'intérêt de cette élimination est de conduire à un système de taille réduite, ce qui donnera une matrice jacobienne plus rapide à inverser à chaque itération de la méthode de Newton. Un inconvénient potentiel est que le système linéaire correspondant peut-être très mal conditionné : nous avons observé des conditionnements proche de 10^{20} ! Il serait intéressant mener une comparaison systématique des deux approches, ce que nous n'avons pu réaliser faute de temps. Pour atténuer le mauvais conditionnement des matrices jacobiennes, nous avons utilisé une technique classique en algèbre linéaire : nous avons équilibré les lignes de la matrice, en pré-multipliant celle-ci par une matrice diagonale contenant la somme des coefficients des lignes (de sorte que celles-ci aient la même norme). Cette modification simple a grandement diminué le nombre de cas où le conditionnement pose un problème. Une alternative utilisée par Hoffmann et al. [46] est de résoudre le système au sens des moindres carrés, ce qui constitue une alternative robuste (mais deux fois plus coûteuse). De plus, ces auteurs résolvent directement le système (3.8).

La méthode numérique la plus utilisée pour réaliser ce calcul est la méthode de Newton, mise en oeuvre dans un grand nombre de codes de géochimie. La thèse de J. Carrayrou [17] contient une comparaison de différentes méthodes de résolution pour le problème de l'équilibre chimique. Sa recommandation est d'utiliser une combinaison de la méthode de Newton avec une méthode de point fixe sur une formulation particulière du système d'équilibre (ce qu'il appelle la méthode des fractions continues positives). Cette combinaison permet de réduire les difficultés dues au manque de convergence globale de la méthode de Newton, si le point initial n'est pas suffisamment proche de la solution (qui est justement ce que l'on cherche à calculer). La difficulté a été aussi rencontrée dans le code Chess [91], qui utilise un « polishing factor » (en fait une étape de recherche linéaire) pour essayer de surmonter les problèmes de convergence globale.

Nous avons choisi d'utiliser une méthode de Newton avec recherche linéaire (implémentée en Matlab par C. T. Kelley, voir [52, 53]) pour assurer cette convergence globale. Il s'agit d'un programme généraliste, auquel nous avons simplement apporté la modification sur la mise à l'échelle indiquée plus haut.

3.1 Exemples

Nous avons validé cette approche sur deux cas test, tirés de la thèse de J. Carrayrou. Le premier est une série d'échanges d'ions, comprenant 6 espèces, et 2 réactions.

Le tableau de Morel est présenté sur la table 3.3. Les composés Mont-K, Mont-Ca et Mont-AL sont des montmorillonites, des complexes formés lors d'un échange avec la roche (voir le paragraphe 4). Dans cet exemple, nous fixons les valeurs de l'itéré ini-

	K^+	Ca^{2+}	Al^{3+}	Mont-K	$\log K$
K^+	1	0	0	0	0
Ca^{2+}	0	1	0	0	0
Al^{3+}	0	0	1	0	0
Mont-K	0	0	0	1	0
Mont-Ca	-2	1	0.2	0.81	
Mont-Al	-3	0	1	3	0.125
Total	-10 mmol	3 mmol	2 mmol	12 mmol	
Initial	10 nmol	variable	variable	11.9 mmol	
Solution	763.7 μ mol	167.4 μ mol	300.5 μ mol	1.237 mmol	

TAB. 3.3: Tableau de Morel, exemple de l'échange d'ions

tial de la méthode de Newton pour deux des composants, et faisons varier les valeurs des autres. La figure 3.1 présente une carte de convergence, c'est-à-dire le nombre d'itérations de la méthode de Newton à partir de ce point initial. Nous avons fait

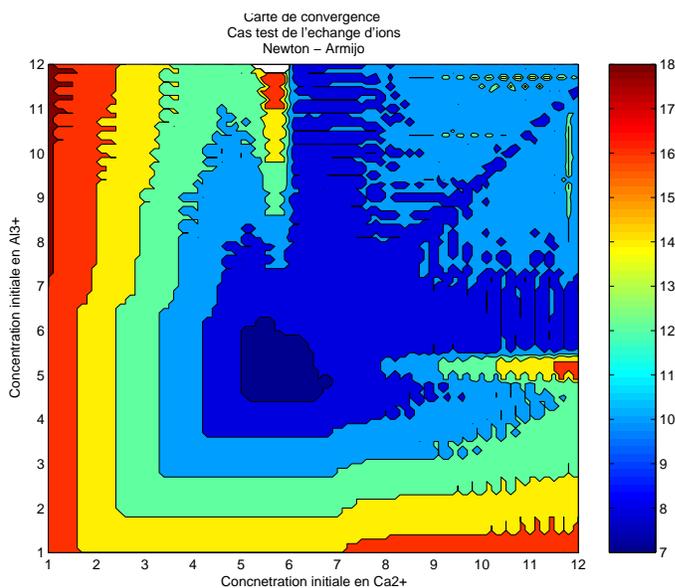


FIG. 3.1: Carte de convergence, exemple de l'échange d'ions

varier les concentrations initiales en Ca^{2+} et Al^{3+} entre 10^{-12} et 10^{-1} (en échelle logarithme). La méthode de Newton, avec recherche linéaire n'a jamais pris plus de 18 itérations, et plutôt entre 6 et 10 pour des concentrations initiales « raisonnables ».

Le deuxième exemple provient de l'article [12], nommé « Acide gallique », et comporte 17 espèces, dont 3 espèces composantes (cet exemple montre bien l'intérêt de la notion de composantes). Le pH est imposé, il n'y a donc que deux concentrations inconnues, à savoir Al^{3+} et H_3L . La figure 3.2 montre la carte de convergence

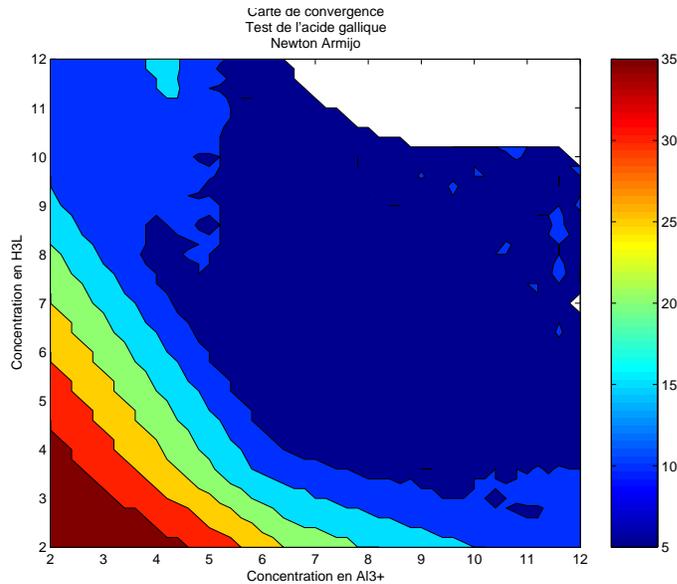


FIG. 3.2: Carte de convergence, exemple de l'acide gallique

pour cet exemple, plus difficile. On voit en effet que si la méthode de Newton démarre avec des valeurs trop lointaines, elle peut ne pas converger. Toutefois, ceci ne se produit que pour des concentrations initiales inférieures à 10^{-10}mol/l , alors que pour la solution on a $[\text{Al}^{3+}] = 2.05 \cdot 10^{-5}\text{mol/l}$ et $[\text{H}_3\text{L}] = 2.59 \cdot 10^{-7}\text{mol/l}$. En dehors de cette région, le nombre d'itérations reste entre 5 et 20, mais peut monter jusqu'à 35.

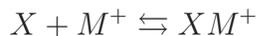
4 Modélisation des réactions de sorption

La sorption est le mécanisme qui décrit l'interaction des espèces dissoutes avec les surfaces solides. Il s'agit essentiellement de mécanismes basés sur des interactions électrostatiques entre l'eau et les composés. On peut distinguer, en première approximation, deux modèles principaux :

- la complexation de surface,
- l'échange d'ions.

La complexation de surface [33], [81] est la liaison entre un ion sorbé et la surface, avec élimination d'une molécule d'eau. On considère alors qu'il y a formation d'un complexe. Par exemple, dans le cas d'un métal monovalent, noté M, on considère qu'il y a formation d'un complexe XM^+ et, avec la réaction chimique analogue à

une réaction de complexation classique



Au niveau microscopique, la modélisation fait intervenir le potentiel de surface, qui contribue une énergie électrique à l'énergie libre de Gibbs. Comme ce potentiel n'est pas directement mesurable, on le relie à la charge de surface, et il existe plusieurs modèles pour cela.

Le modèle d'échange d'ions consiste à considérer que la sorption se passe lorsqu'un métal en solution réagit avec des « sites d'échanges » présents sur la roche, comme dans la réaction [81] :



où l'indice s symbolise le cation fixé sur la surface de l'échangeur. La *capacité d'échange cationique*, ou *CEC* mesure la possibilité limitée de la surface d'accueillir des ions, et donne lieu à une compétition entre ceux-ci, lorsque plusieurs espèces se trouvent en solution.

D'un point de vue pratique, nous nous contenterons d'observer que, quelque soit la modélisation retenue, les réactions de sorption peuvent finalement être considérées comme des réactions chimiques en phase aqueuse, et relèvent de la loi d'action de masse, à condition de définir convenablement la constante d'équilibre de la réaction considérée.

La discussion du paragraphe 3 s'applique donc aux réactions de sorption. Les modifications à apporter sont très simples, et consistent essentiellement à définir les composantes, et les espèces secondaires, aqueuses et sorbées, et à décomposer la matrice stoechiométrique en blocs selon ces définitions. Nous reviendrons au chapitre 4 sur la prise en compte des réactions de sorption, dans le cadre du système couplé avec le transport.

Chapitre 4

Le couplage transport chimie

1 Introduction

Nous abordons maintenant la principale contribution de cette thèse : le couplage entre le transport et la chimie.

La migration d'éléments chimiques à l'état dissous dans les eaux souterraines est un sujet d'actualité qui touche divers domaines (la pollution des nappes souterraines, le stockage de déchets radioactifs). La simulation de l'évolution de ces espèces dans les eaux souterraines nécessite donc un modèle couplé qui prenne en compte à la fois les phénomènes chimiques et le transport des contaminants.

La majorité des méthodes proposées se base sur un couplage pas à pas : on résout itérativement la chimie et le transport. Ce procédé converge, mais parfois au prix d'une réduction draconienne du pas de temps. La résolution du problème de la chimie implique la résolution d'équations différentielles ordinaires (si les réactions sont cinétiques) ou d'équations algébriques non linéaires (si nous supposons que les réactions sont en équilibre local). Le modèle de transport se présente sous la forme d'un système d'équations différentielles (équation d'advection-diffusion). Le système couplé transport-chimie est décrit par un système d'équations algèbro-différentielles aux dérivées partielles, et les méthodes basées sur des variantes robustes de la méthode de Newton ont fait la preuve de leur efficacité dans des domaines variés. La difficulté est bien évidemment la taille du système sous-jacent, dont le nombre d'inconnues est le nombre de points de grille multiplié par le nombre d'espèces chimiques. Une mise en oeuvre naïve de cette méthode (en formant le jacobien du système) est difficile, voire impossible pour des systèmes réalistes, et ne permet pas de conserver les codes de chimie et de transport séparés. Il en résulte que les temps de calculs de ce type de simulation s'élèvent rapidement avec la complexité du problème simulé jusqu'à devenir prohibitifs.

Plusieurs modèles de transport réactif ont été développées suivant deux ap-

proches : l'approche directe (cf. [45], [74], [24], [55], [56]) qui implique la résolution globale de l'équation du transport réactif et l'approche par séparation d'opérateurs qui conduit à résoudre séparément les opérateurs de transport et de chimie (cf. [93], et aussi [50], [60], [74]).

Le problème du transport réactif revient à résoudre à chaque pas de temps un système $f(x) = 0$ où x est l'inconnue, la variable discrète. Lorsque le calcul de $f(x) = 0$ implique un seul code, où le système couplé est résolu en utilisant une variante de la méthode Newton, on parle d'approche globale comme les méthodes GIA (Global Implicit Approach), DSA (Direct Substitution Approach) et s'il implique l'utilisation de n codes, on parle d'approche séquentielle à n pas, où le transport et la chimie sont résolues séparément. On distingue souvent SIA (Sequential Iterative Approach) qui conduit au couplage par itérations entre les termes de transport et de réactions et SNIA (Sequential Non Iterative Approach), où chaque pas du temps consiste en une phase de transport suivie par une phase de réactions. Si l'on alterne transport et chimie, on obtient ce qu'on appelle le Splitting de Strang. Le principe de chacune de ces méthodes est présenté dans la figure (4.1).

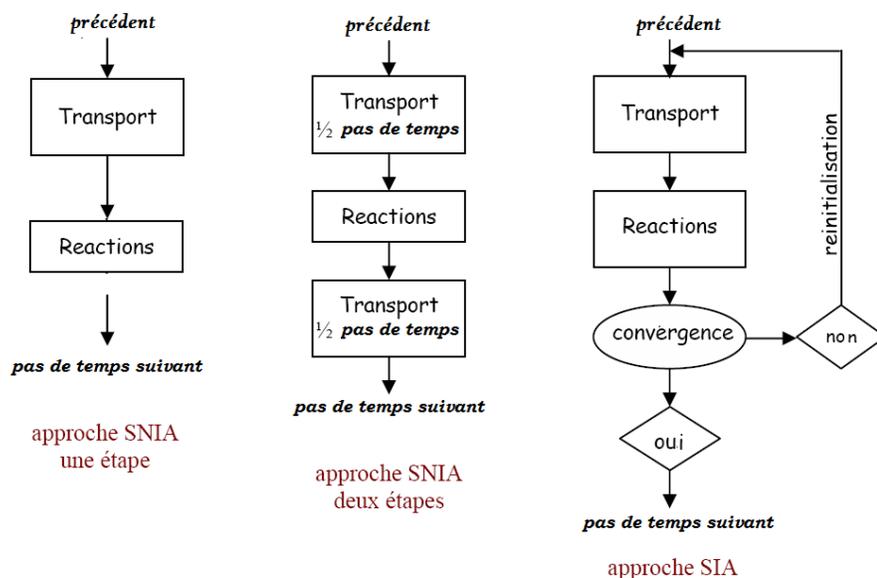


FIG. 4.1: Algorithmes de couplage transport-chimie

Dans la section 2 nous détaillons le modèle choisi, et les méthodes que nous utilisons pour résoudre la partie transport (non-réactif), et l'équilibre chimique. La section 2.3 montre comment nous obtenons le modèle couplé. Les algorithmes de couplage sont le sujet de la section 3, commençant par une discussion des méthodes existantes, alors que notre approche est présentée dans la section 4. Nous donnons des résultats numériques sur le Benchmark MoMaS dans la section 5.

2 Équations du transport réactif

Nous considérons le transport de plusieurs espèces dans un milieu poreux. Les espèces peuvent réagir entre elles et/ou avec la matrice poreuse. Dans cette section nous décrivons les deux modèles, du transport et de la chimie, et les méthodes numériques utilisées pour résoudre les différents sous-systèmes de notre problème couplé.

2.1 Modèle de transport, cf chapitre 2

Le transport des espèces simples soumises à l'advection et la dispersion moléculaire dans un milieu poreux (un domaine $\Omega \subset \mathbf{R}^d$, avec $d = 1, 2$ ou 3), avec une porosité ϕ , et un champ de vitesse Darcy connu u , est régi par l'équation d'advection-diffusion.

$$\phi \frac{\partial c}{\partial t} + L(c) = q, \quad \text{dans } \Omega \quad (4.1)$$

où

$$L(c) = \nabla \cdot (uc) - \nabla \cdot (D\nabla c),$$

est l'opérateur de transport, D est le tenseur de diffusion-dispersion et q est le terme source.

Dans ce travail, nous limitons à un problème monodimensionnel, ainsi nous pouvons écrire l'équation de transport dans un intervalle borné $\Omega =]0, L[$ comme (4.2) :

$$\phi \frac{\partial c}{\partial t} + \frac{\partial}{\partial x} \left(-D \frac{\partial c}{\partial x} + uc \right) = q, \quad 0 < x < L, \quad 0 < t < T \quad (4.2)$$

La condition initiale est $c(x, 0) = c_0(x)$, les conditions limites sont une condition de Dirichlet (la concentration donnée) $c(0, t) = c_d(t)$ au bord à gauche ($x = 0$) et une condition de flux diffusif nul $D \frac{\partial c}{\partial x} = 0$ au bord à droite ($x = L$). Nous pourrions facilement prendre en considération d'autres conditions au bord plus générales.

L'opérateur de transport contient des termes advectifs et diffusifs, il est raisonnable d'utiliser différentes méthodes de discrétisation de temps pour les différents termes. Les termes diffusifs devraient être traités implicitement, et les termes advectifs sont mieux traités explicitement. Plus de détails sur le choix du schéma et la méthode de résolution ont été vu au chapitre 2.

2.2 Équations chimiques, cf chapitre 3

Dans cette section, on décrit le modèle de la chimie. Nous supposons que les réactions sont en équilibre local, ce qui signifie que les phénomènes chimiques se produisent sur une échelle beaucoup plus rapide que des phénomènes de transport.

Nous supposons que nous avons un ensemble d'espèces chimiques $(X_j)_{j=1,\dots,N_e}$ liées par N_R réactions

$$\sum_{j=1}^{N_e} \nu_{ij} X_j \rightleftharpoons 0, \quad i = 1, \dots, N_R,$$

où ν est la matrice stoechiométrique. Selon Morel [64], nous distinguons entre les espèces *composants* et les espèces *secondaires*, en extrayant une sous matrice inversible de la matrice de rang maximal ν .

La loi d'action de masse donne la valeur de l'*activité* en fonction des activités des composants. De même, chaque composant donne lieu à une équation de conservation qui sépare le total des concentrations en composants et espèces secondaires, donc le rôle de la chimie est de séparer les espèces entre celles qui sont en solution, et celles qui ont été absorbées par la roche (nous ne prenons pas en considération la précipitation). Nous notons ainsi

- composantes aqueuses $c_j, j = 1, \dots, N_c,$
- composantes sorbées $s_j, j = 1, \dots, N_s,$
- espèces secondaires aqueuses $x_j, i = 1, \dots, N_x,$
- espèces secondaires sorbées $y_j, i = 1, \dots, N_y.$

Nous avons identifié le nom des espèces avec leurs concentrations, et nous supposons une solution idéale (nous identifions activités et concentrations). Nous supposons également que des espèces secondaires aqueuses peuvent seulement être formées par des composants aqueuses, de sorte que les lois d'action de masse soient écrites sous forme

$$x_i = K_{x_i} \prod_{j=1}^{N_c} c_j^{S_{ij}}, \quad i = 1, \dots, N_x, \quad y_i = K_{y_i} \prod_{j=1}^{N_c} c_j^{A_{ij}} \prod_{j=1}^{N_s} s_j^{B_{ij}}, \quad i = 1, \dots, N_y. \quad (4.3)$$

Les matrices S, A et B sont les matrices stoechiométriques, et les nombres K_{x_i} et K_{y_i} sont des constantes d'équilibre pour chaque réaction.

On a reformulé le problème de la chimie en utilisant comme inconnues principales les logarithmes des concentrations. Ceci a l'avantage supplémentaire que les concentrations sont automatiquement positives. On peut alors récrire la loi d'action de masse (4.3) comme des équations linéaires impliquant les logarithmes des concentrations ($\log u$ est le vecteur dont la i ème composante est $\log u_i$) :

$$\log x = \log K_x + S \log c \quad (4.4)$$

$$\log y = \log K_y + A \log c + B \log s \quad (4.5)$$

Nous exprimons également la conservation de masse pour chaque composante, sous la forme

$$c + S^T x + A^T y = T, \quad s + B^T y = W, \quad (4.6)$$

où T est le vecteur des concentrations totales, et W est relié à la capacité de la matrice de roche d'accepter des ions (voir Dzombak et Morel [34] pour une explication détaillée des phénomènes physiques étant à la base de ces concepts). Nous verrons plus tard que dans le cadre du transport et de la chimie couplés, T est donné par le modèle de transport et fait partie des données du sous problème de la chimie et W est une constante CEC (Cationic Exchange Capacity of the porous matrix) d'échange des ions.

Soient lc , ls , lx , ly respectivement $\log c$, $\log s$, $\log x$, $\log y$. Les équations de conservation de la masse (4.6) peuvent s'écrire encore :

$$\exp(lc) + S^T \exp(lx) + A^T \exp(ly) = T, \quad (4.7)$$

$$\exp(ls) + B^T \exp(ly) = W, \quad (4.8)$$

Pour réduire la taille du système, on élimine les concentrations secondaires, en remplaçant dans (4.7) lx et ly par leurs expressions (4.5) et on résout un système uniquement avec les variables lc et ls , défini par

$$H \begin{pmatrix} lc \\ ls \end{pmatrix} = \begin{pmatrix} \exp(lc) + S^T \exp(\log K_x + Slc) + A^T \exp(\log K_y + Alc + Bls) - T \\ \exp(ls) + B^T \exp(\log K_y + Alc + Bls) - W, \end{pmatrix} \quad (4.9)$$

alors le problème de la chimie est $H \begin{pmatrix} lc \\ ls \end{pmatrix} = 0$. Dans la suite, nous supposons que ce problème a toujours une solution (positive) (c, s) , étant donné les concentrations totales T et W . Le problème d'équilibre chimique est une conséquence de la minimisation de l'énergie de Gibbs ([79], cf chapitre 3, section 2.4).

En conclusion, nous définissons, pour chaque composant sa partie aqueuse (ou mobile) C et sa partie fixe (ou immobile) F , par

$$C = c + S^T x, \quad F = A^T y \quad (4.10)$$

alors $T = C + F$.

La résolution du sous-problème chimique est détaillée dans le chapitre 3, section 3. Pour le problème couplé, une répartition des espèces entre leurs formes fixes et aqueuses est nécessaire plutôt que les différentes concentrations (qui sont nécessaires toujours comme quantités intermédiaires). Il sera commode de condenser le sous-problème chimique par une fonction

$$\begin{aligned} \psi : \mathbf{R}^{Nc} &\rightarrow \mathbf{R}^{Nc} \\ T &\mapsto \psi(T) = A^T y, \end{aligned} \quad (4.11)$$

où y est obtenu en résolvant le problème chimique (4.9) pour lc et ls , étant donné T (et W , que nous prenons comme constante), et en calculant F comme indiqué

ci-dessus. Nous aurons besoin, plus loin, de la Jacobienne de ψ . Son calcul est simple une fois que ψ est calculée. Nous calculons d'abord

$$H' \begin{pmatrix} lc \\ ls \end{pmatrix} = \begin{pmatrix} \text{diag}(\exp(lc)) & 0 \\ 0 & \text{diag}(\exp(ls)) \end{pmatrix} + \begin{pmatrix} S^T & A^T \\ 0 & B^T \end{pmatrix} \begin{pmatrix} \text{diag}(x) & 0 \\ 0 & \text{diag}(y) \end{pmatrix} \begin{pmatrix} S & 0 \\ A & B \end{pmatrix}. \quad (4.12)$$

On a

$$\psi(T) = A^T y = A^T (\exp(A \log c + B \log s + \log Ky)), \quad (4.13)$$

et par un développement de Taylor d'ordre 1, on a :

$$\log(c + \delta c) = \log c + \frac{\delta c}{c} + O(\|\delta c\|^2), \quad \log(s + \delta s) = \log s + \frac{\delta s}{s} + O(\|\delta s\|^2). \quad (4.14)$$

$$\begin{aligned} \psi(T + \delta T) &= A^T \exp \left(A \log c + B \log s + \log Ky + A \frac{\delta c}{c} + B \frac{\delta s}{s} \right) + O(\|\delta c\|^2, \|\delta s\|^2) \\ &= A^T y \exp \left(A \frac{\delta c}{c} + B \frac{\delta s}{s} \right) + O(\|\delta s\|^2, \|\delta c\|^2) \quad (\delta c, \delta s \text{ sont proches de } 0) \\ &= A^T y \left(1 + A \frac{\delta c}{c} + B \frac{\delta s}{s} \right) + O(\|\delta c\|^2, \|\delta s\|^2) \\ &= \psi(T) + A^T y \left(A \frac{\delta c}{c} + B \frac{\delta s}{s} \right) + O(\|\delta c\|^2, \|\delta s\|^2), \end{aligned} \quad (4.15)$$

d'où

$$\psi'(T) \delta T = A^T y \left(A \frac{\delta c}{c} + B \frac{\delta s}{s} \right) = A^T y (A \ B) \begin{pmatrix} \frac{\delta c}{c} \\ \frac{\delta s}{s} \end{pmatrix}. \quad (4.16)$$

et comme $\delta lc = \frac{\delta c}{c}$ et $\delta ls = \frac{\delta s}{s}$, alors

$$\psi'(T) \delta T = A^T y (A \ B) \begin{pmatrix} \delta lc \\ \delta ls \end{pmatrix} = A^T y (A \ B) (H')^{-1} \begin{pmatrix} \delta T \\ \delta W \end{pmatrix}. \quad (4.17)$$

où $\begin{pmatrix} \delta lc \\ \delta ls \end{pmatrix} = (H')^{-1} \begin{pmatrix} \delta T \\ \delta W \end{pmatrix}$ est un petit système linéaire, facile à résoudre.

Pour W constant, on a $\delta W = 0$, d'où finalement

$$\psi'(T) = A^T y (A \ B) (H')^{-1} \begin{pmatrix} I \\ 0 \end{pmatrix}. \quad (4.18)$$

2.3 Couplage transport et chimie

Pour obtenir le modèle couplé entre le transport et la chimie, nous écrivons une équation (4.1) pour chaque espèce et chaque composante (si les espèces sont immobiles, on omet le terme L). Nous supposons que l'opérateur de transport est le même pour toutes les espèces, c.-à-d. $L(c_1, c_2, \dots, c_{N_c}) = (Lc_1, Lc_2, \dots, Lc_{N_c})$.

$$\left\{ \begin{array}{l} \phi \frac{\partial x_i}{\partial t} + L(x_i) = r_i^x, \quad i = 1, \dots, N_x \\ \phi \frac{\partial c_j}{\partial t} + L(c_j) = r_j^c, \quad j = 1, \dots, N_c \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} \phi \frac{\partial y_i}{\partial t} = r_i^y, \quad i = 1, \dots, N_y \\ \phi \frac{\partial s_j}{\partial t} = r_j^s, \quad j = 1, \dots, N_s. \end{array} \right. \quad (4.19)$$

Les termes de r représentent les termes source de réactions, ils sont inconnus. En introduisant le total des concentrations, on peut éliminer ces termes source en utilisant la loi de conservation de la masse. Pour chaque composante j , on forme une combinaison linéaire des équations (4.19) avec le coefficient stoechiométrique approprié, et on obtient le système d'équations suivant qui contient uniquement les concentrations totaux aqueuses et fixées.

$$\left\{ \begin{array}{l} \phi \frac{\partial C_j}{\partial t} + L(C_j) = r_j^c + \sum S_{i,j} r_i^x \\ \phi \frac{\partial W_j}{\partial t} = r_j^s + \sum B_{i,j} r_i^y \\ \phi \frac{\partial F_j}{\partial t} = \sum A_{i,j} r_i^y \end{array} \right.$$

On ajoute la première équation à la dernière, on obtient

$$\phi \frac{\partial C_j}{\partial t} + \phi \frac{\partial F_j}{\partial t} + L(C_j) = r_j^c + \sum S_{i,j} r_i^x + \sum A_{i,j} r_i^y. \quad (4.20)$$

Par utilisation de la loi de conservation de masse, on a $r_j^c + \sum S_{i,j} r_i^x + \sum A_{i,j} r_i^y = 0$ alors

$$\left\{ \begin{array}{l} \phi \frac{\partial C_j}{\partial t} + \phi \frac{\partial F_j}{\partial t} + L(C_j) = 0 \\ \frac{\partial W_j}{\partial t} = 0 \text{ car } W_j \text{ est une constante} \end{array} \right. \quad j = 1, \dots, N_c \quad (4.21)$$

ce qui donne l'équation :

$$\phi \frac{\partial T_j}{\partial t} + L(C_j) = 0 \quad j = 1, \dots, N_c \quad (4.22)$$

où T_j est le total des concentration et C_j est le total des concentrations aqueuses pour le composant j , ces deux quantités sont définies par (4.10) et reliées par l'équation $T_j = C_j + F_j$.

On note aussi L l'opérateur de transport discrétisé. Chaque concentration inconnue dépend de deux indices, l'indice de points de grille, et l'indice des espèces

chimiques. Nous utiliserons une notation inspirée de Matlab. Pour une concentration u_{ij} , où $i \in [1, N_x]$ représente l'indice spatial et $j \in [1, N_c]$ représente l'indice chimique, nous dénoterons par

- $u_{:,j}$ le vecteur de concentrations de l'espèce j à tous les points de la grille ;
- $u_{i,:}$ le vecteur de concentrations de toutes les espèces chimiques au point x_i (techniquement, ceci devrait être un vecteur ligne, mais on néglige cette distinction).

Le problème couplé est obtenu en mettant ensemble l'équation (4.22) et la définition solution de l'opérateur chimique,

$$\begin{cases} \Phi \frac{\partial T_{:,j}}{\partial t} + L(C_{:,j}) = 0, & j = 1, \dots, N_c \\ T_{ij} = C_{ij} + F_{ij}, & i = 1, \dots, N_x, j = 1, \dots, N_c \\ F_{i,:} = \psi(T_{i,:}), & i = 1, \dots, N_x \end{cases} \quad (4.23)$$

où $\Phi = \text{diag}((\phi_i h_i)_{i=1, \dots, N_x})$. Nous discrétisons en temps par un schéma d'Euler implicite avec un pas de temps constant, notant que d'autres stratégies plus sophistiquées sont évidemment possible (en particulier, utiliser un pas de temps variable). On note l'indice de temps en exposant, nous obtenons alors le système couplé suivant, qui est un système non linéaire.

$$\begin{cases} \Phi \frac{T_{:,j}^{n+1} - T_{:,j}^n}{\Delta t} + L(C_{:,j}^{n+1}) = 0 & j = 1, \dots, N_c \\ T_{ij}^{n+1} = C_{ij}^{n+1} + F_{ij}^{n+1} & i = 1, \dots, N_x, j = 1, \dots, N_c \\ F_{i,:}^{n+1} = \psi(T_{i,:}^{n+1}) & i = 1, \dots, N_x \end{cases} \quad (4.24)$$

L'avantage de ce système est qu'on a non seulement éliminé les termes de réaction inconnus, mais aussi que le nombre d'équations de transport a été réduit au nombre de toutes les espèces au nombre de composants (aqueux). Dans la prochaine section, nous discuterons différentes formulations pour le modèle couplé, puis nous présenterons différents algorithmes de couplage, et nous discuterons les méthodes numériques utilisées.

3 Formulations et algorithmes pour le couplage

La formulation du transport réactif vue ci-dessus conduit à un grand système d'équations non linéaires. Pour des problèmes complexes, sa résolution exigera un temps de calcul élevé, donc il est important de choisir une méthode appropriée. Plusieurs approches et formulations sont disponibles.

Grâce à la relation $T = C + F$, il est facile d'éliminer une des 3 variables, et ceci mène à des formulations différentes pour le problème couplé, et cela dépend des variables qui sont gardées dans l'équation de transport.

Les 3 formulations principales (cf Salignac [77]) sont :

- la formulation (TC) où T est la variable principale , C la variable transportée

$$\Phi \frac{\partial T_{:j}}{\partial t} - L(C_{:j}) = 0 \quad (4.25)$$

C'est la formulation utilisée par Erhel et al. [24], pour un type d'algorithme de DAE.

- la formulation (CC) où C est la variable principale , F la variable transportée

$$\Phi \frac{\partial C_{:j}}{\partial t} + \Phi \frac{\partial F_{:j}}{\partial t} - L(C_{:j}) = 0 \quad (4.26)$$

C'est la formulation que nous avons choisie et que nous avons décrite précédemment. L'axe qui a conduit notre choix est la possibilité d'avoir deux codes séparés (de la chimie et du transport) pour pouvoir utiliser des codes existants. En effet, cette formulation nous permet de calculer la quantité des concentrations fixée F par une résolution de la chimie et la quantité mobiles C par une résolution du problème du transport. Cependant cette formulation est moins appropriée pour l'usage des algorithmes de décomposition d'opérateurs, mais dans une méthode globale ceci devrait pas avoir un effet, comme nous itérons jusqu'à la convergence.

- la formulation (TT) où T est la variable principale, F la variable transportée

$$\Phi \frac{\partial T_{:j}}{\partial t} - L(T_{:j}) + L(F_{:j}) = 0. \quad (4.27)$$

Cette formulation ne semble pas satisfaisante, car le transport opère sur les espèces immobiles, et nous l'avons éliminé pour cette raison.

À chaque pas de temps, le système formé par (4.24) (une équation de transport pour chaque composant et un système chimique pour chaque point de grille) forme un grand système non linéaire, dont la taille est le nombre de composants fois le nombre de points de grille.

Ce système a traditionnellement été résolu par l'approche séquentielle en deux-étapes, comme passé en revue ci-dessous (cf. [93]). Cependant, cette méthode a plusieurs inconvénients : elle peut exiger un pas de temps petit pour assurer la convergence, et si l'approche séquentielle utilisée est non-iterative (ce qui est seulement du premier ordre en temps), elle introduit en plus des erreurs de décompositions d'opérateurs. Grâce à sa convergence quadratique, la méthode Newton serait une méthode idéale pour résoudre le système. Cependant, nous rencontrons une difficulté pratique : la méthode de Newton exige la solution d'un système linéaire avec

la matrice Jacobienne à chaque itération. Dans des situations réalistes, il ne sera pas possible de former et factoriser la matrice Jacobienne. Nous verrons dans la prochaine section que cette difficulté peut être surmontée par le recours aux méthodes itératives pour résoudre le système linéaire. Avant de détailler cette étape, nous retournons à une explication de différentes méthodes de couplage.

3.1 Les différentes méthodes de couplage

Il existe différents types d'approche pour le couplage entre le transport et la chimie, on peut distinguer deux catégories : l'approche globale GIA (cf. [45], [74], [24], [55], [56]) et l'approche séquentielle SNIA, SIA, cf. [93], et aussi ([50], [60], [74]).

Méthode séquentielle non-itérative SNIA

Dans cette méthode, un pas de temps consiste en une phase de transport suivie par une phase de réactions utilisant les concentrations du flux transporté. Pour l'équation (4.24), le schéma s'écrit :

$$\begin{cases} \Phi \frac{C_{:,j}^* - C_{:,j}^n}{\Delta t} + L(C_{:,j}^*) = 0 & j = 1, \dots, N_c \\ F_{ij}^* = T_{ij}^* - C_{ij}^* & i = 1, \dots, N_x, j = 1, \dots, N_c \\ \Phi \frac{F_{:,j}^{n+1} - F_{:,j}^*}{\Delta t} = \psi(T_{i,:}^*) & i = 1, \dots, N_x \end{cases} \quad (4.28)$$

Quand les processus physiques sont terminés, les réactions se produisent et modifient la composition chimique de la solution. Son principe est présenté sur la figure (4.1). Cette méthode est attractive en raison de sa simplicité. L'équation du transport peut être résolue espèce par espèce car elles ne sont pas couplées par les réactions. L'équation de chimie peut être résolue de façon locale, car il n'y figure plus aucun terme de transport.

La méthode SNIA présente un avantage au niveau de la validation puisqu'elle couple des codes existants facile à valider séparément. Mais les erreurs dues à la séparation sont importantes ce qui exige un petit pas du temps pour les problèmes raides.

La séparation d'opérateurs de transport et chimie lors de la résolution des équations du transport réactif sous l'hypothèse d'équilibre chimique n'introduit pas, contrairement au cas de la cinétique, des erreurs sur le bilan de masse, en effet, l'opérateur de chimie à l'équilibre conserve la masse, ce qui n'est pas le cas de l'opérateur de la chimie cinétique. Il est possible de réduire l'erreur de splitting en utilisant une méthode centrée en temps nommée "Strang Splitting" (Strang (1968) [83]), qui consiste à effectuer une demi-étape de transport, suivi d'une étape de chimie, suivie d'une demi-étape de transport (voir la figure (4.1)).

Une autre méthode séquentielle permettant de prendre en compte le couplage d'un écoulement multiphasique et la géochimie (transport réactif) a été développée récemment par E. Tillier, A. Michel et L. Trenty [89], [88]. Cette méthode à deux pas s'appelle Rs-Tr. La première étape Rs est une étape de résolution du problème d'écoulement multiphasique et la seconde Tr, est une étape de transport réactif. La méthode Rs-Tr est une méthode de couplage non itérative dans laquelle l'erreur de splitting est corrigée à l'aide d'un terme de pénalisation (le principe est de pénaliser l'équation de transport en ajoutant l'erreur faite sur les concentrations).

Pour s'affranchir de certaines erreurs de séparation d'opérateurs, A. Younes et al. ont utilisé une combinaison entre les méthodes ELLAM avec l'approche DSA ainsi qu'avec l'approche SNIA [94], [95], [37]. Cette combinaison permet une meilleure précision ainsi qu'une meilleure convergence, avec moins d'erreurs de séparation d'opérateurs.

Méthode séquentielle Strang Splitting

Dans cette méthode, un pas du temps consiste une phase chimique est centrée entre deux étapes de transport. Le schéma s'écrit :

$$\left\{ \begin{array}{ll} \Phi \frac{C_{:,j}^* - C_{:,j}^n}{\Delta t/2} + L(C_{:,j}^*) = 0 & j = 1, \dots, N_c \\ F_{ij}^* = T_{ij}^* - C_{ij}^* & i = 1, \dots, N_x, j = 1, \dots, N_c \\ \Phi \frac{F_{:,j}^{**} - F_{:,j}^*}{\Delta t} = \psi(T_{i,:}^*) & i = 1, \dots, N_x \\ C_{ij}^{**} = T_{ij}^{**} - F_{ij}^{**} & i = 1, \dots, N_x, j = 1, \dots, N_c \\ \Phi \frac{C_{:,j}^{n+1} - C_{:,j}^{**}}{\Delta t/2} + L(C_{:,j}^{n+1}) = 0 & j = 1, \dots, N_c \end{array} \right. \quad (4.29)$$

La première équation représente le transport non réactif sur un demi-pas de temps, la 3ème équation représente la chimie sur un pas de temps et la dernière équation représente le transport non réactif sur un demi-pas de temps.

Méthode séquentielle itérative SIA

L'approche séquentielle itérative SIA (sequential iterative approach) permet de corriger l'erreur de séparation d'opérateurs en itérant entre les étapes de transport et la chimie. L'idée de base est d'arriver au pas de temps $n+1$ à une solution entièrement couplée entre les termes de transport et les réactions chimiques. La figure (4.1) présente son principe.

La résolution du problème (4.24) par la méthode SIA revient à résoudre le schéma itératif suivant :

$$\begin{cases} \Phi \frac{C_{:,j}^{n+1,k+1} - C_{:,j}^n}{\Delta t} + \Phi \frac{F_{:,j}^{n+1,k} - F_{:,j}^n}{\Delta t} + L(C_{:,j}^{n+1,k+1}) = 0 & j = 1, \dots, N_c \\ T_{ij}^{n+1,k+1} = C_{ij}^{n+1,k+1} + F_{ij}^{n+1,k} & i = 1, \dots, N_x, j = 1, \dots, N_c \\ F_{i,:}^{n+1,k+1} = \psi(T_{i,:}^{n+1,k+1}) & i = 1, \dots, N_x \end{cases} \quad (4.30)$$

À chaque pas de temps, on résout d'abord une équation de transport pour chaque composant pour $C^{n+1,k+1}$, avec un terme source donné par une concentration fixée et mobile à l'itération précédente. Cette concentration totale dissoute sera ajoutée à la concentration totale fixée $F^{n+1,k}$ (donnée par le sous problème de la chimie à l'itération précédente) pour obtenir $T^{n+1,k+1}$, et celle ci est le total des concentrations utilisé comme une donnée pour résoudre un problème d'équilibre chimique à chaque point de grille.

La convergence est obtenue quand la différence entre les solutions aux itérations k et $k + 1$ est assez petite

$$\frac{\| C^{n+1,k+1} - C^{n+1,k} \|}{\| C^{n+1,k+1} \|} + \frac{\| F^{n+1,k+1} - F^{n+1,k} \|}{\| F^{n+1,k+1} \|} < \epsilon \quad (4.31)$$

Dans la littérature, ceci est connu comme une approche de décomposition d'opérateur, mais c'est plus correctement une méthode de Gauss-Seidel par blocs sur le système couplé, car chaque sous-système est résolu alternativement.

J. Carrayrou et al. [17], [49] ont développé un algorithme associant le schéma séquentiel et une combinaison d'éléments finis discontinus et mixtes hybrides : la partie convective-réactive utilise l'approche SNIA, avec la partie convective résolue par les éléments finis discontinus explicite en temps et la partie dispersive-réactive utilise l'approche SIA, avec la partie dispersive résolue par les éléments finis mixtes hybride implicites en temps. Le couplage entre le transport et la chimie cinétique est réalisée par un schéma itératif symétrique qui assure des erreurs de séparation d'opérateurs nulle sur le bilan de masse et le profil de la concentration.

N. Bouillard et al. [11] utilisent une méthode séquentielle de type Picard accéléré, pour améliorer la méthode de gradient conjugué non linéaire. La méthode demande un petit nombre de dérivées par rapport à ce que ferait une méthode de Newton et peu d'itérations par rapport à un algorithme de Picard standard. Ils ont montré que la méthode de Newton reste rapide sur différents cas.

Approche globale implicite GIA

Il est connu que les méthodes de séparation d'opérateurs mentionnées précédemment ne peuvent pas traiter d'une manière satisfaisante les problèmes difficiles. La

méthode de substitution directe DSA [90], [57] consiste à résoudre les différentes concentrations des composants, par une *substitution* des équations de la chimie dans l'équation du transport. Ceci peut être fait explicitement, cf. [45], en transportant la concentration totale aqueuse C_i pour chaque composante i , ce qui revient à substituer l'équation (4.10) dans l'équation du transport (4.21), on obtient :

$$\begin{cases} \phi \frac{\partial}{\partial t} \left(c_i + \sum_{j=1}^{N_c} S_{ji} x_j \right) + \phi \frac{\partial}{\partial t} \left(\sum_{j=1}^{N_s} A_{ji} y_j \right) + L \left(c_i + \sum_{j=1}^{N_c} S_{ji} x_j \right) = 0, \\ \phi \frac{\partial}{\partial t} \left(s_i + \sum_{j=1}^{N_s} B_{ji} y_j \right) = 0 \end{cases} \quad (4.32)$$

en remplaçant x_j par son expression (4.3), on obtient :

$$\begin{cases} \phi \frac{\partial}{\partial t} \left(c_i + \sum_{j=1}^{N_c} S_{ji} K_{xj} \prod_{k=1}^{N_c} c_k^{S_{jk}} \right) + \phi \frac{\partial}{\partial t} \left(\sum_{j=1}^{N_s} A_{ji} K_{yj} \prod_{k=1}^{N_c} c_k^{A_{jk}} \prod_{k=1}^{N_s} s_k^{B_{jk}} \right) \\ + L \left(c_i + \sum_{j=1}^{N_c} S_{ji} K_{xj} \prod_{k=1}^{N_c} c_k^{S_{jk}} \right) = 0, \\ \phi \frac{\partial}{\partial t} \left(s_i + \sum_{j=1}^{N_s} B_{ji} K_{yj} \prod_{k=1}^{N_c} c_k^{A_{jk}} \prod_{k=1}^{N_s} s_k^{B_{jk}} \right) = 0 \end{cases} \quad (4.33)$$

L'inconvénient de cette formulation est d'introduire une non linéarité dans l'opérateur du transport L .

Pour une autre méthode implicite, cf. [74], les variables C_i dans l'équation du transport seront remplacées par des fonctions \mathcal{C}_i qui expriment la concentration C en fonction des inconnues $\log c$, $\log s$. Cette approche couple fortement la chimie et le transport et parfois le système couplé qui est non linéaire devient compliqué à résoudre, ce qui limite l'application à des systèmes simples (un composant, pas de réactions chimiques entre les différentes espèces, système monodimensionnel...). Cette méthode consiste à résoudre simultanément et globalement les équations de transport et les équations chimiques.

L'avantage principal de l'approche globale est de ne pas introduire d'erreurs de séparation d'opérateurs et son inconvénient est la taille du système à résoudre qui dépend du nombre des espèces et du nombre de points de la discrétisation spatiale.

Cet inconvénient (la taille du système) peut être contourné avec une méthode de réduction développée par Knabner et Krautle [55], [56]. La méthode de réduction est une approche implicite globale qui évite donc les erreurs de séparation d'opérateurs. Cette méthode consiste à réduire le nombre d'équations non linéaires couplées, procédant également par substitution, avec un changement des variables du système

chimique. Ce mécanisme de réduction peut traiter des réactions cinétiques et en équilibre homogènes ou hétérogènes avec les espèces mobiles et immobiles. Cette méthode est plutôt du type implicite, comme celle que nous proposons, car elle ne résoud pas directement les concentrations c et s .

Une implémentation en parallèle (Hammond 2003) [44], [45] peut aussi s'affranchir du problème de la taille du système global. Une autre difficulté est de calculer les dérivées exactes. Il devient parfois difficile de calculer les coefficients de la Jacobienne, une solution est d'utiliser les dérivées numériques, c'est ce qui est fait dans le code PARTAN de Hammond. Ce code approche la matrice Jacobienne par différences finis en utilisant une méthode de Newton classique ou une méthode de Newton sans matrice (JFNK, Jacobien Free Newton Krylov), en approchant des dérivées directionnelles. Cependant cette méthode demande un temps important pour le calcul, et pour accélérer la convergence de la méthode, il devient indispensable d'utiliser à un préconditionneur, le problème de la construction de ce préconditionneur se pose puisqu'on ne possède plus de matrice Jacobienne. Cette difficulté peut être contournée avec des préconditionneurs basés sur la séparation d'opérateurs, on parle parfois d'un préconditionneur physique.

Il est également possible de reformuler le problème comme système un algébro-différentiel (DAE) (cf. C. de Dieuleveult et J. Erhel [24]), obtenu par une approximation des équations différentielles partielles originales par une méthode de lignes. Cette méthode utilise la formulation TC pour le problème couplé, où T est la variable différentielle et C est la variable transportée. Les inconnues du problème de la chimie sont $\log(c)$, $\log(s)$, et p (la concentration due à une précipitation). La concentration totale aqueuse sera exprimée par la fonction \mathcal{C} (en fonction de ces inconnues) par une équation qui s'ajoute au problème couplé, ce dernier s'écrit sous forme de PDAE (Partial Differential Algebraic Equations),

$$\left\{ \begin{array}{l} \Phi \frac{\partial T_i}{\partial t}(x, t) + L(C_i(x, t)) = 0, \quad i, \dots, N_c \\ \varphi(\log(c(x, t)), \log(s(x, t)), p(x, t), T(x, t)) = 0 \\ C_i(x, t) - \mathcal{C}_i(\log(c(x, t)), \log(s(x, t)), p(x, t)) = 0, \quad i, \dots, N_c \end{array} \right.$$

ou encore le système sous forme discrète

$$\left\{ \begin{array}{l} \Phi \frac{\partial T_i}{\partial t}(t) + L(C_i(t)) = 0 \\ \varphi(y(t), T(t)) = 0 \\ C_i(t) - \mathcal{C}_i(y(t)) = 0, \end{array} \right. \quad (4.34)$$

où y est le vecteur des inconnues discrètes $(\log(c), \log(s), p)^T$, et par des transforma-

tions classiques, le système (4.34) s'écrit :

$$\begin{cases} g(t, Y(t), Y'(t)) = 0 \\ Y(t_0) = Y_0 \\ Y'(t_0) = Y'_0. \end{cases}$$

où $Y = (T, C, y)^T$.

Cette méthode a montré son efficacité sur différents cas tests. Cependant un de ses inconvénients est qu'elle exige la formation de la jacobienne, et il devient parfois difficile de calculer ses coefficients, et les dérivés exactes surtout pour un problème de chimie compliqué (précipitation ou réactions cinétiques). De plus, la matrice Jacobienne couple les deux matrices, la matrice Jacobienne de transport, et la matrice Jacobienne de chimie. Le transport et la chimie sont résolus en même temps, ce qui ne permet pas d'utiliser des codes de résolution existants.

4 Approche globale basée sur la méthode Newton–Krylov

La méthode de couplage qu'on a étudié au cours de cette thèse est aussi une méthode directe dont le système à résoudre prend comme inconnues seulement les concentrations totales. C'est différent de l'approche de substitution directe DSA, qui prend comme inconnues les concentrations totale mobiles et fixées, mais également les concentrations pour les espèces composantes qui sont les inconnues principales du problème de la chimie.

Dans notre méthode, le système qu'on doit résoudre à chaque pas de temps exige la résolution du sous-problème chimique à chaque itération de Newton, lors de l'évaluation du résidu. La méthode qu'on a choisie pour résoudre le système non-linéaire est une méthode de Newton-Krylov. Il s'agit d'une méthode de Newton dans laquelle les systèmes linéaires sont résolus par une méthode itérative (le plus souvent GMRES ou BiCGSTAB). Ces méthodes, aussi appelées "Jacobian free", demandent seulement le produit Jacobien-vecteur sans former et stocker la matrice Jacobienne, et le produit peut être approché par différences finies, ce qui est fait dans les travaux de Hammond [45], et aussi [54]. Cependant pour notre méthode, il n'est pas avantageux d'utiliser les dérivées numériques puisqu'elles demandent l'évaluation de la fonction objectif, ce qui exige la résolution du problème non linéaire de la chimie dans notre cas. Pour remédier à ce problème, on a calculé les dérivées directionnelles exactes.

La matrice Jacobienne implique la matrice de transport, aussi bien que la Jacobienne de l'opérateur solution chimique par leur action sur un vecteur. Quoique

cette méthode puisse être plus chère que la méthode précédente DAE [24], son avantage principal est de traiter la chimie comme une boîte-noire, dans le contexte de Newton-Krylov [52]. Il s'agit d'un avantage important, car il permet de conserver les codes de chimie et de transport séparés comme les simulateurs chimiques deviennent de plus en plus sophistiqués.

Nous présentons maintenant notre nouvelle approche. Nous ne substituons pas directement les équations chimiques dans l'équation de transport, mais nous utilisons la fonction ψ présentée précédemment pour représenter l'effet de la chimie. Différentes formulations pourraient être adoptées selon le choix des inconnues. Nous avons choisi d'utiliser les totaux des concentrations fixées et mobiles, comme inconnues principales, et gardons également les concentrations totales (bien qu'elles pourraient facilement être éliminées).

Nous nous rappelons que le système non linéaire à résoudre à chaque pas de temps est :

$$\begin{cases} (\Phi + \Delta t L)C_{:,j}^{n+1} + \Phi F_{:,j}^{n+1} - b_{:,j}^n & = 0, & j = 1, \dots, N_c \\ T^{n+1} - C^{n+1} - F^{n+1} & = 0, \\ F_{i,:}^{n+1} - \psi(T_{i,:}^{n+1}) & = 0, & i = 1, \dots, N_x \end{cases} \quad (4.35)$$

où $b_{:,j}^n = \Phi C_{:,j}^n + \Phi F_{:,j}^n$ est connu.

On note f la fonction

$$f \begin{pmatrix} C \\ T \\ F \end{pmatrix} = \begin{pmatrix} ((\Phi + \Delta t L)C_{:,j} + \Phi F_{:,j} - b_{:,j})_{j \in [1, N_c]} \\ T - C - F \\ (F_{i,:} - \psi(T_{i,:}))_{i \in [1, N_x]} \end{pmatrix}, \quad (4.36)$$

donc on résout $f \begin{pmatrix} C^{n+1} \\ T^{n+1} \\ F^{n+1} \end{pmatrix} = 0$.

Pour résoudre $f(x) = 0$ par la méthode de Newton, on doit à chaque itération nonlinéaire k passer par les étapes suivantes :

1) On évalue le résidu non linéaire f à la nouvelle étape de temps $n+1$. Dans notre problème, l'inconnue $x = (x_1, x_2, x_3)$ est (C, T, F) , l'évaluation du résidu peut être écrite comme

$$f \begin{pmatrix} C^{n+1,k} \\ T^{n+1,k} \\ F^{n+1,k} \end{pmatrix} = \begin{pmatrix} ((\Phi + \Delta t L)C_{:,j}^{n+1,k} + \Phi F_{:,j}^{n+1,k} - b_{:,j}^{n,k})_{j \in [1, N_c]} \\ T^{n+1,k} - C^{n+1,k} - F^{n+1,k} \\ (F_{i,:}^{n+1,k} - \psi(T_{i,:}^{n+1,k}))_{i \in [1, N_x]} \end{pmatrix},$$

2) Pour la méthode de Newton exacte, nous devons former et stocker la matrice Jacobienne, ainsi on forme cette matrice par le calcul des dérivés du residu

$$J_{ij} = \frac{\partial f_i(C^{n+1,k}, T^{n+1,k}, F^{n+1,k})}{\partial x_j} \quad \text{où } i, j \in 1, 2, 3$$

Le Jacobien a une structure par bloc, que nous avons exploitée dans le calcul ci-dessus,

$$J = \begin{pmatrix} (\Phi + \Delta t \mathcal{L}) & 0 & \Phi \\ -I & I & -I \\ 0 & -\Psi'(T) & I \end{pmatrix},$$

où $\mathcal{L} = \text{diag}(L, L, \dots, L)$ et $\Psi'(T)$ est la matrice Jacobienne du problème de la chimie ψ . $\Psi'(T)$ a également une structure de bloc (cependant ce n'est pas un bloc diagonal, car les inconnus sont numérotées par espèces). La structure du Jacobien est illustrée sur la figure (4.2). Nous pouvons clairement voir les différentes parties du Jacobien.

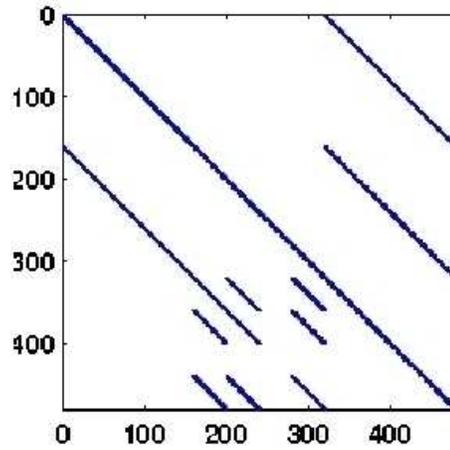


FIG. 4.2: La structure bloc de la matrice Jacobienne

3) On résout le système

$$J \begin{pmatrix} \delta C^k \\ \delta T^k \\ \delta F^k \end{pmatrix} = -f \begin{pmatrix} C^{n+1,k} \\ T^{n+1,k} \\ F^{n+1,k} \end{pmatrix}$$

puis la concentration est mise à jour

$$\begin{pmatrix} C^{n+1,k+1} \\ T^{n+1,k+1} \\ F^{n+1,k+1} \end{pmatrix} = \begin{pmatrix} C^{n+1,k} + \lambda^k \delta C^k \\ T^{n+1,k} + \lambda^k \delta T^k \\ F^{n+1,k} + \lambda^k \delta F^k \end{pmatrix}$$

où λ est déterminée par une recherche linéaire. On doit utiliser une certaine forme de globalisation afin d'assurer la convergence à partir d'un point de départ arbitraire.

Le système linéaire $J\delta x = -f(x)$ est en principe résolu exactement. L'inconvénient principal de la méthode pour les problèmes de grande taille est encore de former, et puis factoriser, la matrice Jacobienne. Une alternative (voir [52], [53], [54] et [45], duquel notre travail s'inspire) est d'utiliser une méthode de Krylov. Il s'agit d'une variante de la méthode de Newton où le système linéaire $J\delta x^k = -f(x^k)$, à chaque itération k de Newton, est résolu par méthode itérative (de type de Krylov).

Détaillons plus précisément la résolution de ce système linéaire. Un premier résidu linéaire r_0 est défini, donné un itéré initial δx_0 (généralement c'est zéro), pour la correction de Newton,

$$r_0 = -f(x) - J\delta x_0$$

Soit j l'indice d'itération de Krylov, la j ème itération minimise $\|J\delta x_j + f(x)\|_2$ par GMRES [73], (ou Bi-CGSTAB [92] ou QMR [39] qui cherchent une solution) dans le sous-espace de Krylov, $K_j = \text{span}\{r_0, Jr_0, J^2r_0, \dots, J^{j-1}r_0\}$, et peut s'écrire sous forme

$$\delta x_j = \delta x_0 + \sum_{i=0}^{j-1} \beta_i J^i r_0$$

où β_i est un scalaire (voir [52], [53], pour les détails sur le calcul et les justifications théoriques).

L'avantage de ce type de méthode est que la formation et le stockage de la matrice Jacobienne n'est pas nécessaire, on doit juste pouvoir calculer le produit de cette matrice avec un vecteur, c'est une dérivée directionnelle, ceci mène aux méthodes Jacobian-Free, où ce produit peut être approchée par des différences finies. Cependant, pour quelques problèmes, il peut être possible de calculer la dérivé directionnelle exactement, et c'est ce qu'on va voir dans la suite, ceci est le cas de notre problème couplé, nous pouvons aussi obtenir le Jacobien du sous-problème chimique.

La formulation donnée précédemment se prête à une implémentation de la méthode de Newton qui permet de garder les deux codes séparés. Le Jacobien a une structure de bloc, les différents codes donnent leur Jacobiens (c'est évidemment plus facile pour le transport que pour la chimie), et ceci reporte au produit Jacobien-vecteur.

Les méthodes Newton-Krylov sont des méthodes inexactes comme le système linéaire n'est pas résolu exactement. Le critère d'arrêt pour les itérations linéaires est de la forme

$$\|J\delta x + f(x^k)\| \leq \eta_k \|f(x^k)\| \quad (4.37)$$

Le choix du terme η_k doit vérifier les deux points suivants :

- garder la convergence locale des méthodes de Newton ;
- éviter la résolution trop précise du système linéaire, ie. prendre moins d'itérations linéaires lorsqu'on est loin de la solution non-linéaire.

Le premier point tendra à exiger une petite valeur pour η_k , alors que le deuxième tend évidemment à rendre η_k plus grand. Un choix possible (cf. [53]) est donné par l'expression :

$$\eta_k = \gamma \|f(x_k)\|^2 / \|f(x_{k-1})\|^2, \quad (4.38)$$

où $\gamma \in [0, 1]$ est un paramètre (la valeur suggérée dans [52] est $\gamma = 0.9$).

Le produit Jacobien par vecteur peut être approché par une dérivée directionnelle, comme on peut facilement le voir par un développement de Taylor de f .

Etant donné un point x , et un vecteur v , l'approximation est

$$Jv = f'(x)v \approx \frac{f(x + \varepsilon v) - f(x)}{\varepsilon}. \quad (4.39)$$

Ceci signifie que, pour les fonctions complexes f , il n'est pas nécessaire de calculer la matrice Jacobienne, au coût de l'évaluation supplémentaire de la fonction lui-même. Il s'avère, cependant, que dans notre cas, cette façon de procéder n'est pas avantageuse. En effet, la partie la plus chère de l'évaluation de f est la solution du problème chimique à chaque point de grille. D'une part, nous avons vu que le calcul du Jacobien de ψ est moins coûteux que le calcul de la fonction ψ elle-même, car il implique seulement la solution d'un système linéaire (voir l'équation (4.18)), tandis que le calcul de ψ exige la résolution d'un système non linéaire. Utilisant le calcul dans la section 2.2, nous pouvons calculer l'action du Jacobien sur un vecteur comme

$$J \begin{pmatrix} v_C \\ v_T \\ v_F \end{pmatrix} = \begin{pmatrix} ((\Phi + \Delta L)v_{C:,j} + \Phi v_{F:,j})_{j \in [1, N_c]} \\ -v_C + v_T - v_F \\ (v_{F i,:} - \psi'(T_{i,:}, v_{T i,:}))_{i \in [1, n_x]} \end{pmatrix}. \quad (4.40)$$

Le premier terme se calcule espèce par espèce, et le troisième se calcule point par point, en utilisant (4.18).

La méthode Jacobian-Free Newton-Krylov est plus robuste que la méthode de Newton basée sur une résolution directe du système linéaire, dans le sens qu'elle est moins sensible à la taille du système.

5 Résultats numériques

5.1 Échange d'Ions : exemple EX11 de PhreeqC

L'exemple suivant du transport advectif en présence d'échange d'ions est adopté comme un premier cas test pour la comparaison des deux approches. L'exemple est présenté dans la documentation de PHREEQC-2 [68] comme exemple 11. Le problème de simulation monodimensionnel décrit une expérience dans une colonne où la composition chimique contenant un échange d'ions est simulée. Initialement, la colonne contient une solution de sodium-potassium-nitrate en équilibre. On injecte à gauche de la colonne trois volumes de pore d'une solution de chlorure de calcium. Le calcium, le potassium, et le sodium réagissent à l'équilibre avec l'échangeur. Le chlore joue le rôle d'un traceur. Les paramètres d'écoulement et de transport utilisés pour cet exemple sont présentés dans le tableau (4.1), et les concentrations initiales aqueuses sont représentées dans le tableau (4.2). La capacité d'échange des ions est 1.1 mmol/l.

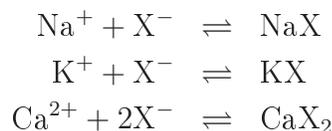
Vitesse de Darcy	$2.78 \cdot 10^{-6}$ m/s
Coefficient de diffusion	$5.56 \cdot 10^{-9}$ m ² /s
Longueur de la colonne	0.08 m
Pas d'espace	0.0002 m
Durée de l'expérience	1 jour
Pas du temps	720 s

TAB. 4.1: paramètres physiques

Composantes	C_{init}	C_{inflow}
Ca	0	$0.6 \cdot 10^{-3}$
Cl	0	$1.2 \cdot 10^{-3}$
K	$2.0 \cdot 10^{-4}$	0
Na	$1.0 \cdot 10^{-3}$	0

TAB. 4.2: Concentrations aqueuses

Les réactions chimiques pour cet exemple peuvent s'écrire :



KX, NaX, CaX₂, sont les complexes formés avec la roche. X indique le site d'échange avec la charge -1. La description chimique (loi d'action de masse et de conservation) peut être représentée dans le tableau Morel suivant :

	Ca ²⁺	Na ⁺	Cl ⁻	K ⁺	X	log K
CaX ₂	1	0	0	0	2	3.4576
NaX	0	1	0	0	1	0
KX	0	0	0	1	1	0.7
Totale	T_{Ca}	T_{Na}	T_{Cl}	T_K	W	

où T_{Ca} , T_{Na} , T_{Cl} , T_K sont les concentrations totales aqueuses et W est la concentration totale fixée (dans cet exemple W est une constante qui vaut $1.1 \cdot 10^{-3}$).

Comparaison avec Phreeqc

Nous appliquons la méthode séquentielle et notre méthode globale sur le cas test décrit précédemment et comparons les résultats obtenus par les deux méthodes. Les résultats obtenus par les deux approches sont analogues, ils sont représentés sur la même figure (4.3) à gauche, et ils sont semblables aux résultats de Phreeqc [68], représentés sur la figure (4.3) à droite. Les deux figures montrent des courbes d'évolution de la concentration des diverses espèces à l'extrémité de la colonne, en fonction de temps. Les ions absorbés de potassium et de sodium sont successivement remplacés par le calcium. Le chlore Cl^- ne réagit pas avec les autres espèces, c'est un traceur.

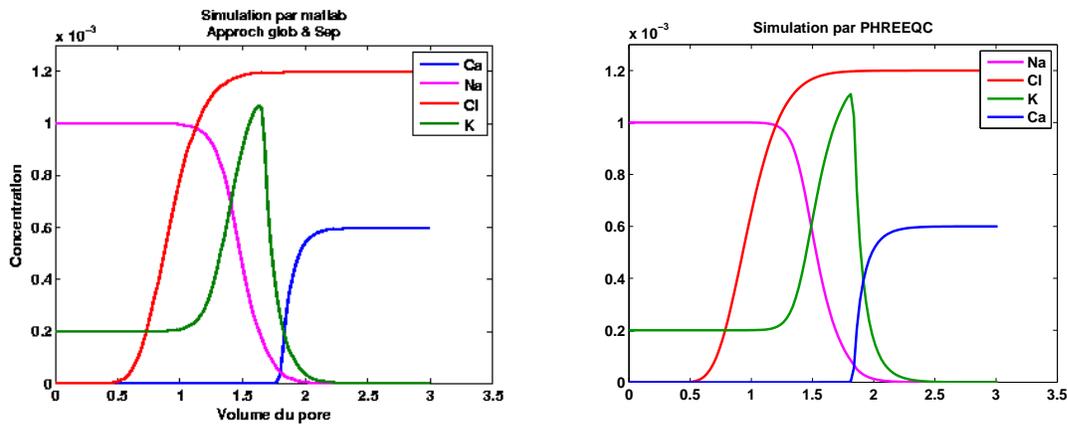


FIG. 4.3: évolution des concentrations

Comparaison du temps CPU pour les trois méthodes

Nous sommes intéressés également par la comparaison de temps CPU pour l'approche séquentielle itérative et non itérative et l'approche globale. Nous traçons dans la figure (4.4) la courbe du temps CPU demandé par chaque méthode en fonction du nombre de noeuds de la grille. Nous remarquons que l'approche globale demande moins de temps CPU que les méthodes itératives séquentielles et plus de temps CPU que l'approche séquentielle non itérative et ceci pour un problème 1D avec une chimie simple, tel que celui considéré ici. Entre deux pas du temps, le calcul par approche séquentielle itérative demande entre 20 et 27 itérations, et ceci pour un maillage grossier. Le nombre de ces itérations augmente avec le nombre de noeuds de la grille. Entre deux pas de temps, le calcul par approche globale demande moins de 6 itérations non linéaires et ce nombre d'itérations de Newton reste stable même

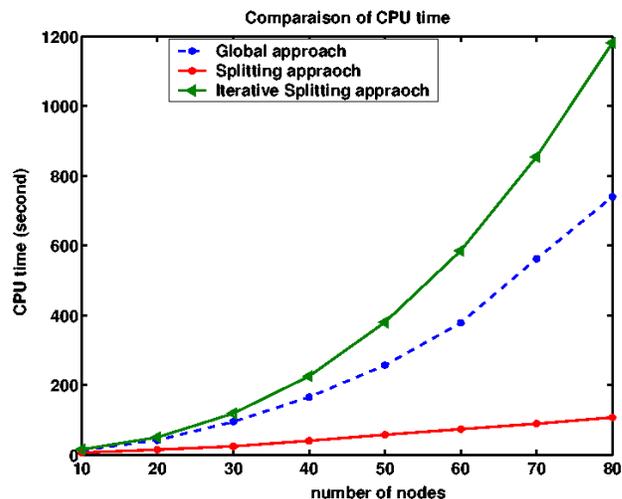


FIG. 4.4: Calcul du temps CPU pour les 3 méthodes

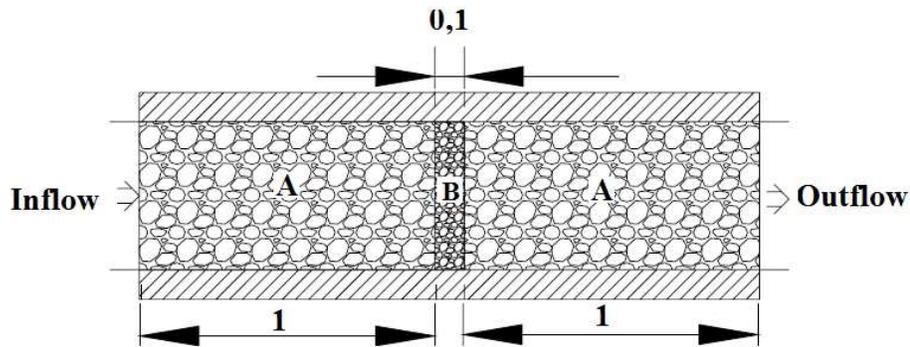
lorsqu'on raffine le maillage, mais le nombre d'itérations de Krylov augmentent avec la taille de la grille. Ainsi la plus grande partie de temps pris par la méthode globale est due aux itérations linéaires. Pour éviter ce problème et améliorer cette méthode, une solution possible serait d'utiliser un bon preconditionner afin d'accélérer la convergence et réduire le temps CPU de l'approche globale. Faute du temps nous n'avons pas pu aborder ce sujet.

5.2 Benchmark GDR MoMaS : 1D easy

Le benchmark est décrit dans [18], [19]. L'objectif de cet exercice est de comparer les méthodes numériques utilisées pour résoudre un problème de transport réactif dans les milieux poreux. Selon les intérêts du GDR MoMaS, le problème du transport réactif devrait être représentatif des problèmes rencontrés dans des simulations de déchets nucléaires : géochimie, géologie hydraulique, méthodes numériques, mathématiques appliquées. Les problèmes proposés ici sont établis sur les mêmes concepts mathématiques que des problèmes hydrogéochimiques réels, mais leur description a été simplifiée. La difficulté pour établir ce benchmark était également de fournir un problème simple suffisant sans perte de difficultés mathématiques et numériques.

a) L'exercice

Géométrie :1D



Le domaine est un milieu poreux hétérogène, composé de deux milieux A et B. Le milieu A est caractérisé par une grande perméabilité (\mathcal{K}), une faible porosité (ω), dispersivité (α_L) et réactivité (W) par rapport au milieu B.

$$A : \quad \mathcal{K} = 0.01, \quad \omega = 0.25, \quad W = 1, \quad \alpha_L = 10^{-2}$$

$$B : \quad \mathcal{K} = 10^{-5}, \quad \omega = 0.5, \quad W = 10, \quad \alpha_L = 6 \cdot 10^{-2}.$$

La vitesse de pore est donnée par : $\omega \cdot u = 5.5 \cdot 10^{-3} \text{ (L.T}^{-1}\text{)}$ dans tout le domaine.

Afin d'être proche des cas réalistes, la frontière et les conditions initiales sont exprimées pour des concentrations totales. On impose une concentration totale au bord à gauche (Inflow) du domaine et un gradient de concentration nul au bord à droite (Outflow), et un calcul de la chimie sépare cette concentration imposées en concentration totale dissoute et concentration totale fixée.

La période d'injection correspond aux concentrations spécifiques (Tab.4.3) injectées au bord Inflow durant la période de temps $[0 \text{ T}, 5000 \text{ T}]$, et la période de lessivage correspond aux concentrations (Tab.4.3) injectées après le temps 5000 T .

	X_1	X_2	X_3	X_4	S	K
c_1	0	-1	0	0	0	10^{-12}
c_2	0	1	1	0	0	1
c_3	0	-1	0	1	0	1
c_4	0	-4	1	3	0	0.1
c_5	0	4	3	1	0	10^{+35}
cs_1	0	3	1	0	1	10^{+6}
cs_2	0	-3	0	1	2	10^{-1}
Total initial	T_1	T_2	T_3	T_4	W	
milieu A	0	-2	0	2	1	
milieu B	0	-2	0	2	10	
Injection ($t \in [0, 5000]$)	0.3	0.3	0.3	0		
Lessivage ($t \in [5000, ..]$)	0	-2	0	2		

TAB. 4.3: Tableau de Morel, équilibre chimique

Ce tableau représente les coefficients stoechiométriques pour les lois d'action de masse et l'équation de conservation. Les lois d'action de masse sont données par ligne et celles de conservation par colonne. On notera d'une part la très grande variation des ordres de grandeur des constantes d'équilibre, d'une part la présence de grands coefficients stoechiométrique, qui rendent le problème très non linéaire.

Loi d'action de masse

$$\begin{aligned}
 c_1 &= 10^{-12} X_2^{-1} \\
 c_2 &= 1 X_2 X_3 \\
 c_3 &= 1 X_2^{-1} X_4 \\
 c_4 &= 0.1 X_2^{-4} X_3 X_4^3 \\
 c_5 &= 10^{+35} X_2^4 X_3^3 X_4
 \end{aligned}$$

Loi de conservation de masse

$$\begin{aligned}
 T_1 &= X_1 \\
 T_2 &= X_2 - c_1 + c_2 - c_3 - 4c_4 + 4c_5 + 3cs_1 - 3cs_2 \\
 T_3 &= X_3 + c_2 + c_4 + 3c_5 + cs_1 \\
 T_4 &= X_4 + c_3 + 3c_4 + c_5 + cs_2 \\
 W &= s + cs_1 + 2cs_2
 \end{aligned}$$

Le domaine est initialement en équilibre local avec le composant S, en présence des composants aqueux X_2 et X_4 . Pendant l'injection, le composant X_4 sera enlevé. Le composant X_1 est un traceur parfait, X_2 réagira avec X_3 , avec le composant S et avec ce qui reste de X_4 . Pendant la période de lessivage, X_1 et X_3 seront enlevés. X_2 et X_4 vont réagir avec la surface S et avec ce qui reste du composant X_3 .

b) Les résultats numériques

On applique l'approche globale sur cet exemple, nous commençons par des résultats montrant l'évolution des espèces composantes à diverses instants, et utilisant plusieurs résolutions spatiales (et temporelles).

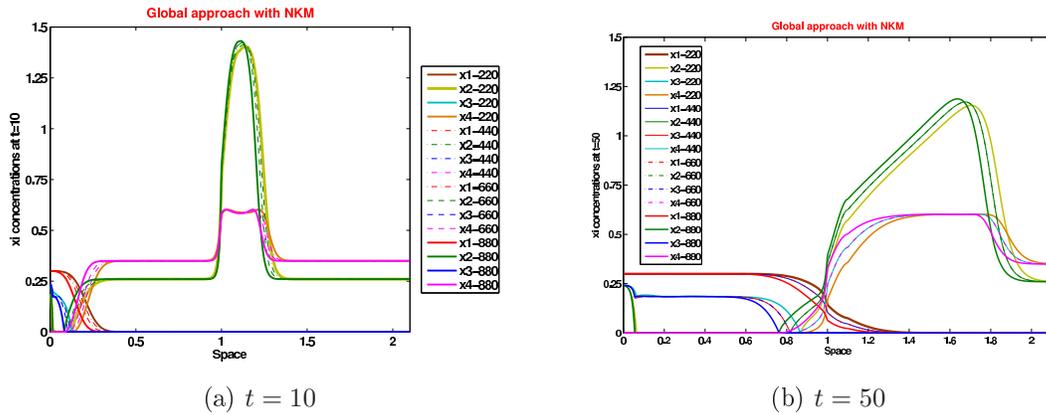


FIG. 4.5: Concentration des composants X1 et X4

On a utilisé un maillage uniforme avec 220, 440, 660 et 880 points, et pour chaque cas, on a choisi comme pas du temps, 0.9 fois le maximum autorisé par la condition CFL. On montre sur la figure (4.6) les courbes d'élutions, à gauche pour t de 0 à 400 (figure 4.6a) et à droite pour t de 4900 à 5300 (figure 4.6b).

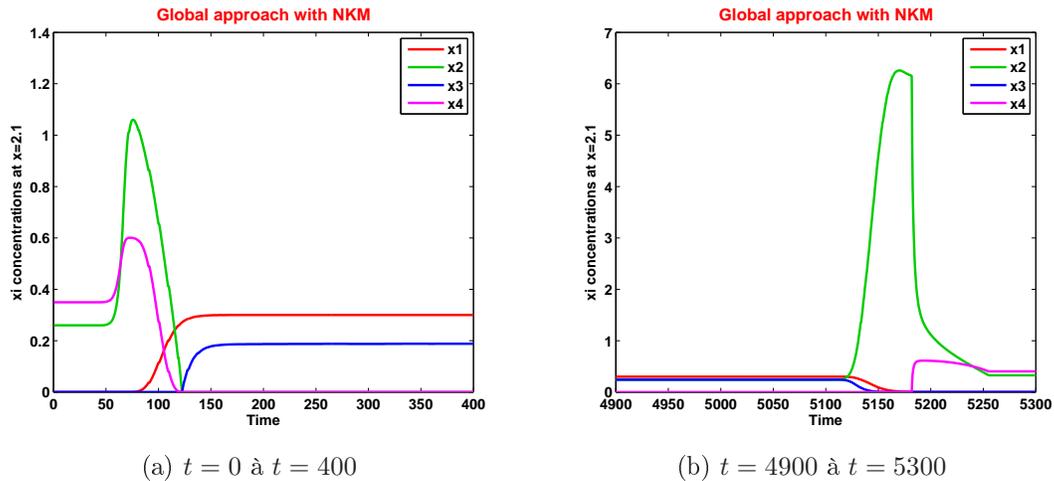


FIG. 4.6: Concentration des composants X1 et X4 au bord de la colonne ($x = 2.1$) en fonction du temps

Nous montrons aussi les résultats demandés dans la définition du Benchmark. La plupart ont été obtenues avec le maillage 220 points, qui peut ne pas être suffisant, comme on le montre ci-dessous. Il n'a pas encore été possible d'obtenir des résultats avec une résolution plus fine de maille pendant des périodes plus longues.

La figure (4.7a) (courbe d'élution pour le total des concentrations dissoute du composant X3) montre les oscillations qui ont été observées par d'autres groupes participant au benchmark. On trouve les oscillations aussi sur la figure (4.7b). Nous pensons que les oscillations ne sont pas d'origine chimique ou physique, mais plutôt d'origine numérique et c'est ce qui est montré par V. Lagneau et J. van der Lee [57] : ces oscillations sont conséquence de la discrétisation en espace pour le système d'équilibre chimique associé aux espèces instables, de forte réactivité.

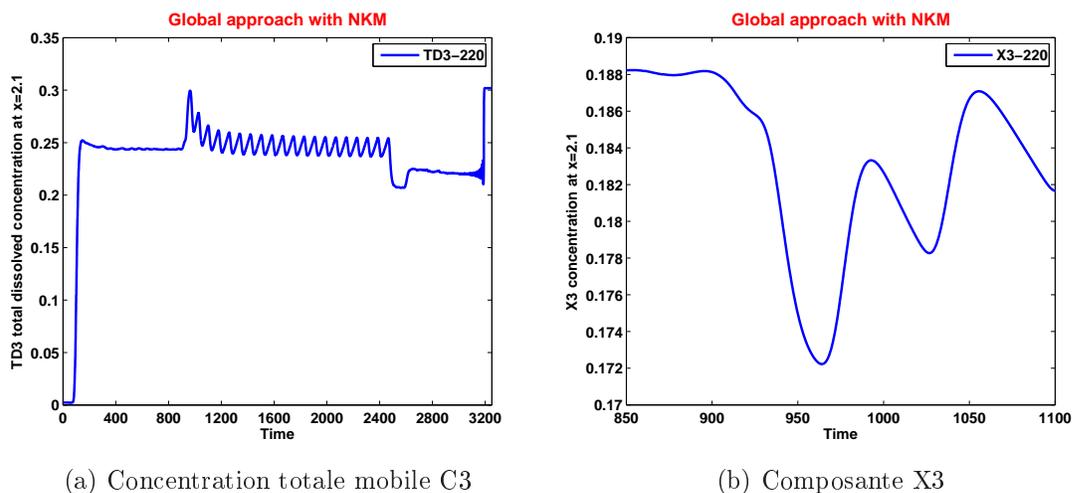


FIG. 4.7: Courbes d'élutions (concentrations en $x = 2.1$ en fonction du temps)

Les figures (4.8a) et (4.8b) montrent l'influence du maillage. Avec un maillage grossier (220 points), les résultats ne sont pas satisfaisants, et avec le maillage 660 points les résultats sont meilleurs. Cependant, ceci peut ne pas être suffisant pour éliminer les oscillations montrées sur figure (4.7), une difficulté est que raffiner le maillage peut ne pas être assez. Comme le problème non linéaire devient plus difficile, il peut être nécessaire d'augmenter le nombre maximum d'itérations pour assurer la convergence de la méthode de Newton-Krylov.

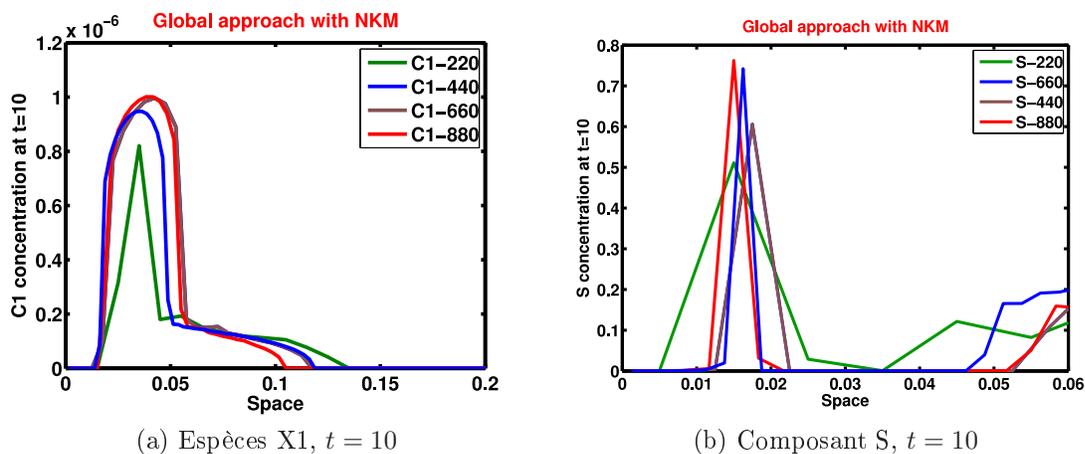


FIG. 4.8: Profils de concentrations

Performance de la méthode

Le benchmark est conçu pour être un cas test difficile pour des méthodes numériques, et c'est en effet le cas. En moyenne, plus de 20 itérations de Newton sont nécessaires à chaque pas de temps, et entre 15 et 40 itérations de GMRES sont nécessaires à chaque itération non linéaire.

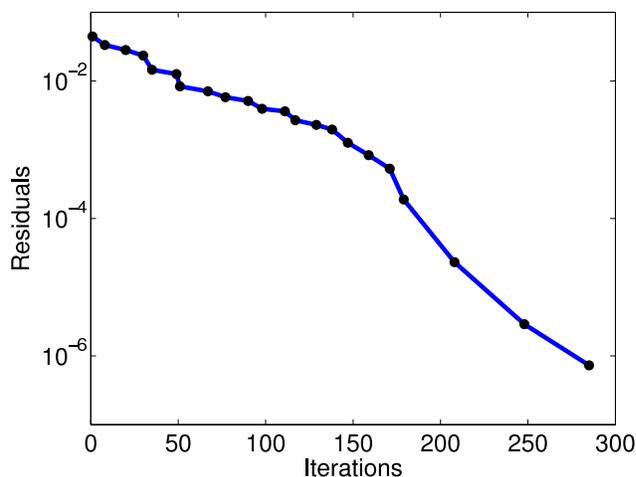


FIG. 4.9: Itérations

La figure (4.9) montre un pas de temps typique : la courbe continue montre le nombre cumulatif d'itération de GMRES (alternativement, le nombre de produits de matrice vecteur), et les points représentent les itérations non linéaires.

Itérations	Maillage 220		Maillage 440		Maillage 660	
	Non lin	Lin	Non lin	Lin	Non lin	Lin
Step 1	16	176	22	285	25	268
Step 2	15	214	22	328	22	342
Step 3	25	272	25	246	25	244
Step 4	18	306	20	417	25	363
Step 5	12	301	16	368	25	620
Step 101	18	377	22	682	25	814
Step 102	19	427	21	684	25	824
Step 103	25	494	18	551	25	636
Step 104	18	476	21	579	25	562
Step 105	15	426	22	583	21	741
Step 201	25	387	25	529	25	625
Step 202	19	437	25	689	24	802
Step 203	14	334	22	634	22	752
Step 204	16	408	20	559	24	829
Step 205	16	420	16	572	22	696

TAB. 4.4: statistique des itérations linéaires et non linéaires

Le tableau (4.4) représente une statistique pour trois intervalles de temps différents, et trois maillages différents. Pour chaque pas de temps, on représente le nombre de pas de Newton et le nombre cumulé des itérations linéaires.

Le nombre d'itérations non linéaires semble dépendre faiblement de la résolution du maillage, et le nombre d'itérations linéaires augmente avec la résolution du maillage.

Chapitre 5

Theoretical and numerical study for flow in a porous medium with intersecting fractures

Cette partie sera l'objet d'une publication à soumettre sur une étude théorique du modèle d'intersection des fractures. Les résultats numériques ont été publiés dans l'article [5]. Ce chapitre est écrit en anglais, on n'a pas traduit le papier pour garder la version originale. Dans ce chapitre, on démontre l'existence et l'unicité de la solution pour le modèle d'écoulement qui prend en compte l'intersection des fractures, et on représente les résultats numériques sur des cas tests académiques.

Abstract

This paper concerns flow in a porous medium with one or more faults or fractures. In the model to be presented, the fractures are treated as interfaces between neighboring subdomains. A domain decomposition method, in which the transmission conditions are non local, is formulated. In an earlier article [61] the model for a domain with a single fracture was defined, existence and uniqueness of the solution were shown, error estimates were obtained and numerical results were shown for a simple two dimensional problem. In the paper [5], the case of intersecting fractures was treated by imposing continuity of the pressure and continuity of the flux at the intersection of the fractures. Numerical results for the model were shown. Here in this paper, we demonstrate the existence and uniqueness of the solution of the model with intersecting fractures.

1 Introduction

A porous medium is in general highly heterogenous. This is due in part to the presence of fractures. Taking into account the heterogeneties is one of the main

difficulties in the modeling of flow in the subsurface. The permeability in fractures may differ greatly from that of the surrounding medium. The fractures can act as hydraulic conductors, providing easy pathways for fluid flow, or as barriers that prevent flow across them. In addition, the width of the fracture is very small in comparison to the size of the domain and this makes numerical modeling of flow in the medium with a fracture quite difficult. Even if the width of the fracture is small, the flow in the fracture can't be neglected, and it's important to model the fractures and take them into account in flow simulation because they can contribute to the capacity of the ground to transport water and pollutants, according to whether they are more or less permeable than the surrounding medium.

One can distinguish two types of fractures : the small fractures which appear in great number in the medium forming a regular network of cracks whose precise localizations are impossible to determine. They are often taken into account using a double porosity model [29], [30] or by stochastic modeling [27], [26]. Other fractures are larger fractures which represent geological discontinuities and these can be localised in space. It is the case of the larger fractures which interests us here.

Among the methods which deal with the problem of the larger fractures in a numerical model, most obvious are those which are based on local mesh refinement around the fracture. Another approach presented in [2] consists in treating more permeable fracture as an interface through which the pressure is continuous but flow is discontinuous. This model is generalized to the case of less permeable fractures in [61], coupling on the interface the jumps of the pressure and normal velocity. A model for Forchheimer fractures is given in [40]. The case of intersecting fractures was taken in [2], and numerical tests in 2D were given, then in [5] numerical tests in 3D were carried out.

In this work, we will present some analysis and theoretical study to validate the model of intersecting fractures. This work also involves the testing of a preconditioner for the iterative method used to solve the problem.

To model the pressure field, and to simulate correctly the fluid flow, we discretize the flow equation in each subdomain and in each fracture using the mixed finite element method, so that the velocity is calculated at the same time as the pressure and the scheme is conservative.

2 Flow in a porous medium : problem with fractures

Let Ω be a convex domain in \mathbb{R}^n , $n = 2$ or 3 , and $\partial\Omega$ the boundary of Ω . We consider an incompressible, single phase fluid in Ω , so the flow, in the absence of gravity, is governed by two laws :

- the law of mass conservation

$$\operatorname{div} \vec{u} = q \quad \text{on} \quad \Omega, \tag{5.1}$$

with \vec{u} the Darcy velocity and q the source term,

– and Darcy's law ;

$$\vec{u} = -K\vec{\nabla}p \quad \text{on } \Omega, \quad (5.2)$$

where p is the pressure, and K is the permeability tensor. We suppose that K is diagonal and satisfies

$$0 < K_{\min}\|x\|_0^2 \leq (Kx, x) \leq K_{\max}\|x\|_0^2, \quad \forall x \neq 0.$$

We add the boundary condition :

$$p = p_d \quad \text{on } \Gamma = \partial\Omega \quad (5.3)$$

where $\vec{\nu}$ is a unit exterior normal vector field on Γ .

For simplicity of exposition we have supposed throughout the following, except in the section with the numerical results, that the boundary condition is a Dirichlet boundary condition. However the model can easily be defined for the case of mixed boundary conditions.

We first present a description of the model given in [61] for a domain with a single fracture. Suppose that the fracture Ω_f is a subdomain of Ω and that there is a

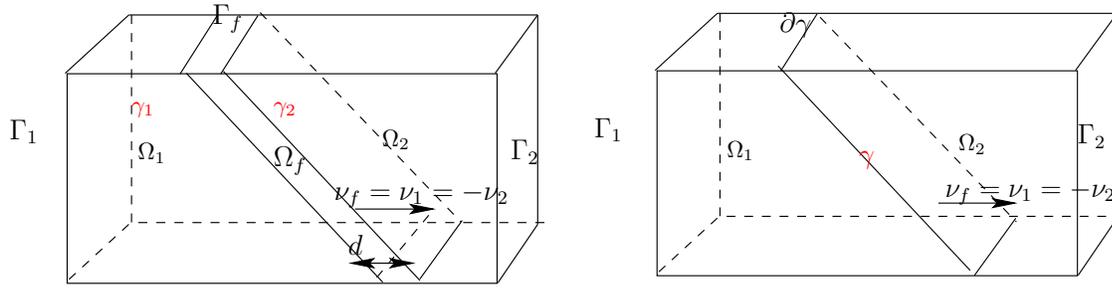


FIG. 5.1: Left : the domain Ω with the fracture Ω_f . Right : the subdomains Ω_1 and Ω_2 separated by the fracture considered as an interface γ .

hyperplane γ with unit vector ν_f normal to γ such that

$$\Omega_f = \left\{ x \in \Omega : x = s + r\vec{\nu}_f \quad \text{for some } s \in \gamma \quad \text{and some } r \text{ in the interval } \left(-\frac{d(s)}{2}, \frac{d(s)}{2}\right) \right\},$$

where $d(s)$ is the thickness of the fracture at $s \in \gamma$. We assume also that $\overline{\Omega}_f$ separates Ω into two connected subdomains Ω_1 and Ω_2 :

$$\Omega \setminus \overline{\Omega}_f = \Omega_1 \cup \Omega_2, \quad \Omega_1 \cap \Omega_2 = \emptyset.$$

Let Γ_i be the part of the boundary of Ω_i coinciding with the boundary of Ω :

$$\Gamma_i = \partial\Omega_i \cap \partial\Omega, \quad i = 1, 2, f,$$

and we denote by γ_i part of the boundary of Ω_i coinciding with the boundary of Ω_f :

$$\gamma_i = \partial\Omega_i \cap \partial\Omega_f \cap \Omega, \quad i = 1, 2.$$

We denote by p_i, u_i, K_i and q_i respectively the restrictions of p, u, K and q to Ω_i and p_{di} the restriction of p_d to $\Gamma_i, i = 1, 2, f$. We can rewrite problem ((5.1),..., (5.3)) as a transmission problem :

$$\begin{aligned}
 (\mathcal{P}_f) \quad & \begin{aligned}
 \operatorname{div} \vec{u}_i &= q_i & \text{in } \Omega_i, & \quad i = 1, 2, f \\
 \vec{u}_i &= -K_i \vec{\nabla} p_i & \text{in } \Omega_i, & \quad i = 1, 2, f \\
 p_i &= p_{di} & \text{on } \Gamma_i, & \quad i = 1, 2, f \\
 p_i &= p_f & \text{on } \gamma_i, & \quad i = 1, 2 \\
 \vec{u}_i \cdot \vec{\nu}_i &= \vec{u}_f \cdot \vec{\nu}_f & \text{on } \gamma_i, & \quad i = 1, 2.
 \end{aligned}
 \end{aligned}$$

In the model presented here, the fracture is treated as an interface between the domains Ω_1 and Ω_2 . By averaging along the line segments $[s - d(s)\vec{\nu}_f, s + d(s)\vec{\nu}_f]$, we obtain the interface model for the fracture problem as in [61],

$$\begin{aligned}
 (\mathcal{P}_\gamma) \quad & \begin{aligned}
 \operatorname{div} \vec{u}_i &= q_i & \text{in } \Omega_i, & \quad i = 1, 2 \\
 \vec{u}_i &= -K_i \vec{\nabla} p_i & \text{in } \Omega_i, & \quad i = 1, 2 \\
 \operatorname{div}_\tau \vec{u}_\gamma &= q_\gamma + (\vec{u}_1 \cdot \vec{\nu}_1 + \vec{u}_2 \cdot \vec{\nu}_2) & \text{on } \gamma \\
 \vec{u}_\gamma &= -K_{f,\tau} d \vec{\nabla}_\tau p_\gamma & \text{on } \gamma \\
 p_i &= p_{di} & \text{on } \Gamma_i, & \quad i = 1, 2 \\
 -\xi \vec{u}_i \cdot \vec{\nu}_i + \frac{2K_{f,\nu}}{d} p_i &= \frac{2K_{f,\nu}}{d} p_\gamma - (1 - \xi) \vec{u}_j \cdot \vec{\nu}_j & \text{on } \gamma, & \quad i, j \in \{(1, 2), (2, 1)\} \\
 p_\gamma &= p_{d\gamma} & \text{on } \partial\gamma,
 \end{aligned}
 \end{aligned}$$

where

$$\vec{u}_\gamma = \int_{-d/2}^{d/2} \vec{u}_{f,\tau} dr, \quad p_\gamma = \frac{1}{d} \int_{-d/2}^{d/2} p_f dr, \quad q_\gamma = \int_{-d/2}^{d/2} q_f dr, \quad p_{d\gamma} = \int_{-d/2}^{d/2} p_{df} dr,$$

and ξ is a quadrature parameter, $\xi > \frac{1}{2}$, $\vec{u}_{f,\tau}$ and $K_{f,\tau}$ are the tangential components of \vec{u}_f and K_f respectively, and $K_{f,\nu}$ is the normal component of K_f .

For the study of intersecting fractures, we consider a simple model in which the domain Ω , is subdivided into 3 subdomains $\Omega_i, i = 1, 2, 3$, naturally separated by fractures considered as interfaces see Figure 5.2.

We use the following notation :

$$I = \{1, 2, 3\} \quad S_i = \{j \in I, i \neq j\}, \quad S = \{(i, j) : i \in I, j \in S_i \text{ and } i < j\}.$$

Then

$$\bar{\Omega} = \bigcup_{i=1}^3 \bar{\Omega}_i, \quad \Omega_i \cap \Omega_j = \emptyset \quad \text{and} \quad \bar{\Omega}_i \cap \bar{\Omega}_j = \bar{\gamma}_{ij} \quad \text{for } i \in I = \{1, 2, 3\}, j \in S_i,$$

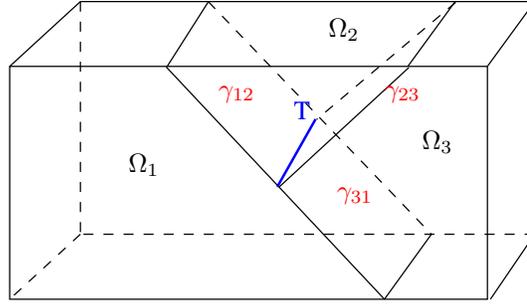


FIG. 5.2: Geometry of the domain Ω

and we let

$$\begin{aligned} \Gamma_i &= \partial\Omega \cap \partial\Omega_i, \quad T = \bar{\gamma}_{12} \cap \bar{\gamma}_{23} \cap \bar{\gamma}_{31}, \\ \gamma_{ij} &= (\bar{\gamma}_{ij} \cap \Omega) \setminus T \quad \text{for } i \in I, j \in S_i, \quad \gamma = \bigcup_{(i,j) \in S} \gamma_{ij}. \end{aligned}$$

For each subdomain $\Omega_i, i \in I$, we have again the usual flow equations (mass conservation and Darcy's law) :

$$\operatorname{div}(\vec{u}_i) = f_i \quad \text{on } \Omega_i \quad (5.4)$$

$$\vec{u}_i = -K_i \vec{\nabla} p_i \quad \text{on } \Omega_i, \quad (5.5)$$

where p_i, \vec{u}_i, f_i and K_i are respectively the pressure, the Darcy velocity, the source term, and the permeability tensor in Ω_i .

The conditions on the exterior boundaries of Ω_i can be written as

$$p_i = p_{di} \quad \text{on } \Gamma_i. \quad (5.6)$$

On the interfaces $\gamma_{i,j}, i \in I, j \in S_i$, we impose as in [61]

$$p_i = p_{i,j} + \frac{d_{i,j}}{2K_{i,j}^\nu} (\xi \vec{u}_i \vec{\nu}_i - (1 - \xi) \vec{u}_j \vec{\nu}_j) \quad \text{in } \gamma_{i,j}, \quad (5.7)$$

where $p_{i,j}$ is the pressure in the fracture $\gamma_{i,j}$, $K_{i,j}^\nu$ is the permeability of the fracture in the direction normal to the fracture, $d_{i,j}$ is the width of the fracture and the parameter ξ is greater than $1/2$.

Remark : Equation (5.7) takes into account a pressure jump across the fracture. It is obtained by integrating the equation $\vec{u} \cdot \vec{\nu} = -K_{f,\nu} d \vec{\nabla} p \cdot \vec{\nu}$, (the normal component in Darcy's law) across the fracture using a quadrature rule depending on ξ , cf [61]. Note that, as the width $d_{i,j}$ of the fracture is quite small, if in addition the fracture is highly permeable (in the direction normal to the fracture) i. e. if $K_{i,j}^\nu$ is sufficiently large, then the second term of the right-hand side of equation (5.7)

is negligible, and equation (5.7) may be replaced by an equation expressing the the continuity of the pressure across the fracture :

$$p_i = p_{i,j} = p_j \quad \text{in} \quad \gamma_{i,j}. \quad (5.8)$$

This is the model which was studied in [1] and [2], and that we have used for the numerical examples in this article.

For the second transmission condition on the interfaces $\gamma_{i,j}$, instead of continuity of flow, one has, as in [1] and [61], a system of equations which model flow in the fractures. The discontinuity of the flux at an interface representing a fracture between two neighboring subdomains yields a source term in the equations describing the flow in that fracture :

$$\operatorname{div}_\gamma(\vec{u}_{i,j}) = q_{i,j} + (\vec{u}_i \cdot \vec{\nu}_i + \vec{u}_j \cdot \vec{\nu}_j) \quad \text{in} \quad \gamma_{i,j} \quad (5.9)$$

$$\vec{u}_{i,j} = -d_{i,j} K_{i,j}^\tau \vec{\nabla}_\gamma p_{i,j} \quad \text{in} \quad \gamma_{i,j} \quad (5.10)$$

$$p_{i,j} = p_{d\gamma_{i,j}} \quad \text{in} \quad \partial\gamma_{i,j} \cap \partial\Omega \quad (5.11)$$

where $K_{i,j}^\tau$ is the tangential component of the permeability tensor of the fracture $\gamma_{i,j}$, $p_{d\gamma_{i,j}}$ is the pressure given on its Dirichlet boundary, $q_{i,j}$ is the source term on the fracture and $\vec{\nu}_{i,j}$ is the outward-pointing, unit, normal vector field on $\partial\gamma_{i,j}$. The two operators $\operatorname{div}_\gamma$ and ∇_γ are, respectively, the surface divergence and the surface gradient on γ .

On the segment T , itself an n-2 dimensional interface at which the fractures meet, the continuity of pressure and the continuity of flux are imposed :

$$p_{i,j} = p_T \quad \text{in} \quad T, \quad \forall i, j \in S_i; \quad i \neq j \quad (5.12)$$

$$\vec{u}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{u}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{u}_{1,3} \cdot \vec{\nu}_{1,3} = 0 \quad \text{in} \quad T, \quad (5.13)$$

where p_T is the pressure on T , and $\vec{\nu}_{i,j}$ is the exterior normal vector to $\gamma_{i,j}$ on T .

3 The weak formulation of the problem

To give the weak formulation of the problem, we need to define two spaces : W a space for the velocity and M a space for the pressure.

$$W = \left\{ \begin{aligned} u = (\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_{1,2}, \vec{u}_{2,3}, \vec{u}_{1,3}) \in H(\operatorname{div}, \Omega_1) \times H(\operatorname{div}, \Omega_2) \times H(\operatorname{div}, \Omega_3) \\ \times H(\operatorname{div}_{\gamma_{1,2}}, \gamma_{1,2}) \times H(\operatorname{div}_{\gamma_{2,3}}, \gamma_{2,3}) \times H(\operatorname{div}_{\gamma_{1,3}}, \gamma_{1,3}) : \forall (i, j) \in S, \\ \vec{u}_i \cdot \vec{\nu}_i + \vec{u}_j \cdot \vec{\nu}_j \in L^2(\gamma_{i,j}) \quad \text{and} \quad (\vec{u}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{u}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{u}_{1,3} \cdot \vec{\nu}_{1,3}) \in L^2(T) \end{aligned} \right\}$$

$$M = \left\{ \begin{aligned} p = (p_1, p_2, p_3, p_{1,2}, p_{2,3}, p_{1,3}, p_T) \in L^2(\Omega_1) \times L^2(\Omega_2) \times L^2(\Omega_3) \times L^2(\gamma_{1,2}) \times \\ L^2(\gamma_{2,3}) \times L^2(\gamma_{1,3}) \times L^2(T) \end{aligned} \right\}.$$

with the norms

$$\begin{aligned} \|u\|_W^2 &= \sum_{i \in I} (\|\vec{u}_i\|_{0,\Omega_i}^2 + \|\operatorname{div} \vec{u}_i\|_{0,\Omega_i}^2) + \sum_{(i,j) \in S} (\|\vec{u}_{i,j}\|_{0,\gamma_{ij}}^2 + \|\operatorname{div}_{\gamma_{ij}}(\vec{u}_{i,j})\|_{0,\gamma_{ij}}^2) \\ &+ \sum_{(i,j) \in S} \|\vec{u}_i \cdot \vec{\nu}_i + \vec{u}_j \cdot \vec{\nu}_j\|_{0,\gamma_{ij}}^2 + \|\vec{u}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{u}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{u}_{1,3} \cdot \vec{\nu}_{1,3}\|_{0,T}^2 \\ \|p\|_M^2 &= \sum_{i \in I} \|p_i\|_{0,\Omega_i}^2 + \sum_{(i,j) \in S} \|p_{i,j}\|_{0,\gamma_{ij}}^2 + \|p_T\|_{0,T}^2. \end{aligned}$$

The spaces W et M are Hilbert spaces.

To obtain the mixed formulation we define two bilinear forms $\alpha : W \times W \longrightarrow \mathbb{R}$ and $\beta : W \times M \longrightarrow \mathbb{R}$ by

$$\begin{aligned} \alpha(\vec{u}, \vec{v}) &= \sum_{i \in I} \int_{\Omega_i} K_i^{-1} \vec{u}_i \cdot \vec{v}_i + \sum_{(i,j) \in S} \int_{\gamma_{i,j}} (d_{i,j} K_{i,j}^T)^{-1} \vec{u}_{i,j} \cdot \vec{v}_{i,j} \\ &+ \sum_{(i,j) \in S} \int_{\gamma_{i,j}} \frac{d_{i,j}}{2K_{i,j}^n} [(\xi \vec{u}_i \cdot \vec{\nu}_i - (1 - \xi) \vec{u}_j \cdot \vec{\nu}_j) \vec{v}_i \cdot \vec{\nu}_i + (\xi \vec{u}_j \cdot \vec{\nu}_j - (1 - \xi) \vec{u}_i \cdot \vec{\nu}_i) \vec{v}_j \cdot \vec{\nu}_j] \\ \beta(\vec{u}, r) &= \sum_{i \in I} \int_{\Omega_i} \operatorname{div} \vec{u}_i r_i + \sum_{(i,j) \in S} \int_{\gamma_{i,j}} (\operatorname{div}_{\gamma_{i,j}} \vec{u}_{i,j} - (\vec{u}_i \cdot \vec{\nu}_i + \vec{u}_j \cdot \vec{\nu}_j)) r_{i,j} \\ &+ \int_T (\vec{u}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{u}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{u}_{1,3} \cdot \vec{\nu}_{1,3}) r_T, \end{aligned}$$

and let the linear forms $L_f : M \longrightarrow \mathbb{R}$ and $L_d : W \longrightarrow \mathbb{R}$ be defined by

$$\begin{aligned} L_f(r) &= \sum_{i \in I} \int_{\Omega_i} f_i r_i + \sum_{(i,j) \in S} \int_{\gamma_{i,j}} f_{i,j} r_{i,j} \\ L_d(v) &= \sum_{i \in I} \int_{\Gamma_i} p_{di} \vec{v}_i \cdot \vec{\nu}_i + \sum_{(i,j) \in S} \int_{\partial\gamma_{i,j} \cap \Gamma} p_{d\gamma_{ij}} \vec{v}_{i,j} \cdot \vec{\nu}_{i,j}. \end{aligned}$$

With these definitions of spaces and forms, the weak formulation of the problem (5.4) through (5.13) can be written as follows :

$$\begin{aligned} &\text{Find } (\vec{u}, p) \in W \times M \text{ such that} \\ (\mathcal{P}) \quad &\alpha(\vec{u}, \vec{v}) - \beta(\vec{v}, p) = -L_d(\vec{v}), \quad \forall \vec{v} \in W \\ &\beta(\vec{u}, r) = L_f(r), \quad \forall r \in M. \end{aligned}$$

4 Existence and uniqueness of the solution

Before showing the existence and uniqueness of the solution of problem (\mathcal{P}) , we first recall some theorems concerning traces and normal traces on subsets of the boundary of a domain. (The subset of interest in the following section will be the subset T on the boundary of each of the fractures $\gamma_{i,j}$ i. e. the space \mathcal{O} considered below will correspond to any one of the fractures $\gamma_{i,j}$ and the subspace Γ of its

boundary to T . In this section we have supposed that the boundary condition on Γ_i , $i = 1, \dots, 3$, is a homogenous dirichlet condition).

Suppose that \mathcal{O} is an open subset of \mathbb{R}^n , with $n = 1, 2$, with boundary $\partial\mathcal{O}$.

Theorem 1. (well known in [22]) *The mapping $r \rightarrow r|_{\Gamma}$ defined a priori for continuous functions r on $\bar{\mathcal{O}}$ can be extended to a continuous linear mapping γ called the trace mapping from $H^1(\mathcal{O})$ into $L^2(\partial\mathcal{O})$. The Sobolev space $H^{\frac{1}{2}}(\partial\mathcal{O})$ may be characterized as the image of the trace mapping γ , and an equivalent norm on this space is $\|\phi\|_{\frac{1}{2},\partial\mathcal{O}} = \inf_{\{r \in H^1(\mathcal{O}) : \gamma r = \phi\}} \|r\|_{1,\mathcal{O}}$.*

Theorem 2. (well known in [41]) *The mapping $\vec{w} \rightarrow \vec{w} \cdot \nu$ defined a priori from $(H^1(\mathcal{O}))^n$ into $L^2(\partial\mathcal{O})$ can be extended to a continuous linear mapping γ_{ν} from $H(\text{div}, \mathcal{O})$ onto $H^{-\frac{1}{2}}(\partial\mathcal{O})$. The space $H^{-\frac{1}{2}}(\partial\mathcal{O})$, the dual of $H^{\frac{1}{2}}(\partial\mathcal{O})$, can also be characterized as the image of the normal trace mapping ν and an equivalent norm on this space is given by $\|\mu\|_{-\frac{1}{2},\partial\mathcal{O}} = \inf_{\{\vec{w} \in H(\text{div}, \mathcal{O}) : \nu \vec{w} = \mu\}} \|\vec{w}\|_{\text{div}, \mathcal{O}}$.*

Green's theorem : For $\vec{w} \in H(\text{div}, \mathcal{O})$ and $r \in H^1(\mathcal{O})$ we have

$$\int_{\mathcal{O}} \text{div} \vec{w} r \, dx + \int_{\mathcal{O}} \vec{w} \cdot \nabla r \, dx = \int_{\partial\mathcal{O}} \vec{w} \cdot \vec{n} r \, ds,$$

where $\int_{\partial\mathcal{O}} \vec{w} \cdot \vec{n} r \, ds = \langle \nu \vec{w}, \gamma r \rangle_{H^{-\frac{1}{2}}(\partial\mathcal{O}), H^{\frac{1}{2}}(\partial\mathcal{O})}$.

Suppose that Γ is a proper subset of $\partial\mathcal{O}$. The space $H_{00}^{\frac{1}{2}}(\Gamma)$ is defined ?? to be the space of L^2 functions φ on Γ whose extension $\tilde{\varphi}$ by zero to all of $\partial\mathcal{O}$ is an $H^{\frac{1}{2}}$ function on $\partial\mathcal{O}$:

$H_{00}^{\frac{1}{2}}(\Gamma) = \{\varphi \in L^2(\Gamma) : \tilde{\varphi} \in H^{\frac{1}{2}}(\partial\mathcal{O})\}$ with norm $\|\varphi\|_{H_{00}^{\frac{1}{2}}(\Gamma)} = \|\tilde{\varphi}\|_{H^{\frac{1}{2}}(\partial\mathcal{O})}$ and $H_{00}^{-\frac{1}{2}}(\Gamma)$ is its dual. Finally, we will use also the spaces $H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$ and $H_{0,\partial\mathcal{O}\setminus\Gamma}^{\frac{1}{2}}(\partial\mathcal{O})$ defined by

$$\begin{aligned} H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O}) &= \{\rho \in H^1(\mathcal{O}) : \gamma \rho|_{(\partial\mathcal{O}\setminus\Gamma)} = 0\} \\ H_{0,\partial\mathcal{O}\setminus\Gamma}^{\frac{1}{2}}(\partial\mathcal{O}) &= \{\eta \in H^{\frac{1}{2}}(\partial\mathcal{O}) : \eta = \text{tr}(\psi) \text{ for some } \psi \in H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})\}. \end{aligned}$$

Lemma 4.1. *The function $\phi \in L^2(\Gamma)$ belongs to $H_{00}^{\frac{1}{2}}(\Gamma)$ if and only if $\tilde{\phi}$, its extension by 0 to all of $\partial\mathcal{O}$ belongs to $H_{0,\partial\mathcal{O}\setminus\Gamma}^{\frac{1}{2}}(\partial\mathcal{O})$; i.e.*

$$\phi \in H_{00}^{\frac{1}{2}}(\Gamma) \quad \Longleftrightarrow \quad \tilde{\phi} \in H_{0,\partial\mathcal{O}\setminus\Gamma}^{\frac{1}{2}}(\partial\mathcal{O}).$$

Proof. Let $\varphi \in H_{00}^{\frac{1}{2}}(\Gamma)$, then $\tilde{\varphi} \in H^{\frac{1}{2}}(\partial\mathcal{O})$ (by definition of space $H_{00}^{\frac{1}{2}}(\Gamma)$), so there exists $\hat{\varphi} \in H^1(\mathcal{O})$ such that $\tilde{\varphi} = \gamma \hat{\varphi}$ (by definition of the space $H^{\frac{1}{2}}(\partial\mathcal{O})$). But as $\gamma \hat{\varphi}|_{(\partial\mathcal{O}\setminus\Gamma)} = \tilde{\varphi}|_{(\partial\mathcal{O}\setminus\Gamma)} = 0$ we have $\hat{\varphi} \in H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$, and $\tilde{\varphi} \in H_{0,\partial\mathcal{O}\setminus\Gamma}^{\frac{1}{2}}(\partial\mathcal{O})$.

Conversely suppose that $\psi \in H_{0,\partial\mathcal{O}\setminus\Gamma}^{\frac{1}{2}}(\partial\mathcal{O})$, so there exists $\check{\psi} \in H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$ such that $\psi = \gamma(\check{\psi})$. But then $\psi|_{(\partial\mathcal{O}\setminus\Gamma)} = (\gamma\check{\psi})|_{(\partial\mathcal{O}\setminus\Gamma)} = 0$, and the extension by 0 of $\psi|_{\Gamma}$ to all of $\partial\mathcal{O}$ is just ψ itself which belongs to $H^{\frac{1}{2}}(\partial\mathcal{O})$. Thus $\psi|_{\Gamma} \in H_{00}^{\frac{1}{2}}(\Gamma)$. \square

Lemma 4.2. *Suppose that $r \in L^2(\mathcal{O})$ and $\lambda \in L^2(\Gamma)$. Then the problem*

$$\int_{\mathcal{O}} \nabla\varphi \cdot \nabla\psi \, dx = \int_{\mathcal{O}} r \psi \, dx - \int_{\Gamma} \lambda \psi \, ds \quad \forall \psi \in H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O}) \quad (5.14)$$

has a unique solution φ in $H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$ and if $\vec{v} = -\nabla\varphi$ then $\vec{v} \in H(\text{div}, \mathcal{O})$ and $\vec{v} \cdot \vec{\nu} \in L^2(\Gamma)$.

Proof. That a unique solution φ exists is guaranteed by the theorem of Lax-Milgram as the bilinear form $a(\varphi, \psi) = \int_{\mathcal{O}} \nabla\varphi \cdot \nabla\psi \, dx$ is elliptic on $H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O}) \times H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$ (Poincarés lemma) and the linear form $\mathcal{L}(\psi) = \int_{\mathcal{O}} r \psi \, dx - \int_{\Gamma} \lambda \psi \, ds$ is continuous on $H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$ (the trace theorem).

Let $\varphi \in H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$ be the solution of (5.14) for a given $r \in L^2(\mathcal{O})$ and $\lambda \in L^2(\Gamma)$, and let $\vec{v} = -\nabla\varphi$. Then $\vec{v} \in (L^2(\mathcal{O}))^n$ since $\varphi \in H_{0,\partial\mathcal{O}\setminus\Gamma}^1(\mathcal{O})$. To see that $\text{div } \vec{v} \in L^2(\mathcal{O})$ it suffices to note that $\text{div } \vec{v}$ is a distribution on \mathcal{O} defined by : for any $\xi \in \mathcal{C}_0^\infty(\mathcal{O})$

$$\begin{aligned} \langle \text{div } \vec{v}, \xi \rangle_{\mathcal{D}(\mathcal{O}), \mathcal{C}_0^\infty(\mathcal{O})} &= \int_{\mathcal{O}} -\vec{v} \cdot \nabla\xi \, dx \\ &= \int_{\mathcal{O}} \nabla\varphi \cdot \nabla\xi \, dx \\ &= \int_{\mathcal{O}} r \xi \, dx - \int_{\Gamma} \lambda \xi \, ds \\ &= \int_{\mathcal{O}} r \xi \, dx \quad (\text{because } \xi \in \mathcal{C}_0^\infty(\mathcal{O})) \\ &= \langle r, \xi \rangle_{\mathcal{D}(\mathcal{O}), \mathcal{C}_0^\infty(\mathcal{O})} . \end{aligned}$$

Then as $r \in L^2(\mathcal{O})$ it follows that $\text{div } \vec{v} \in L^2(\mathcal{O})$, and $\vec{v} \in H(\text{div}, \mathcal{O})$.

As $\vec{v} \in H(\text{div}, \mathcal{O})$ it follows that $\vec{v} \cdot \vec{\nu} \in H^{-\frac{1}{2}}(\partial\mathcal{O})$, and $\vec{v} \cdot \vec{\nu}$ is an element of $H_{00}^{-\frac{1}{2}}(\Gamma)$ defined by :

for any element $\psi \in H_{00}^{-\frac{1}{2}}(\Gamma)$,

$$\begin{aligned}
 \langle \vec{v} \cdot \vec{v}, \psi \rangle_{H_{00}^{-\frac{1}{2}}(\Gamma), H_{00}^{\frac{1}{2}}(\Gamma)} &= \langle \vec{v} \cdot \vec{v}, \tilde{\psi} \rangle_{H^{-\frac{1}{2}}(\partial\mathcal{O}), H^{\frac{1}{2}}(\partial\mathcal{O})} \\
 &= \int_{\mathcal{O}} \operatorname{div} \vec{v} \tilde{\psi} \, dx + \int_{\mathcal{O}} \vec{v} \cdot \nabla \tilde{\psi} \, dx \\
 &= \int_{\mathcal{O}} r \tilde{\psi} \, dx + \int_{\mathcal{O}} -\nabla \phi \cdot \nabla \tilde{\psi} \, dx \\
 &= \int_{\Gamma} \lambda \tilde{\psi} \, ds \\
 &= \int_{\Gamma} \lambda \psi \, ds \\
 &= \langle \lambda, \psi \rangle_{H_{00}^{-\frac{1}{2}}(\Gamma), H_{00}^{\frac{1}{2}}(\Gamma)},
 \end{aligned} \tag{5.15}$$

where for $\psi \in H_{00}^{-\frac{1}{2}}(\Gamma)$, $\tilde{\psi} \in H_{0, \partial\mathcal{O} \setminus \Gamma}^{\frac{1}{2}}(\partial\mathcal{O})$ is its extension by 0 to $\partial\mathcal{O}$ and $\check{\psi} \in H_{0, \partial\mathcal{O} \setminus \Gamma}^1(\mathcal{O})$ is an extension of $\tilde{\psi}$ to \mathcal{O} , and where Green's formula has been used for the second equality and lemma 4.1 for the fourth equality to guarantee that $\check{\psi}$ is in $H_{0, \partial\mathcal{O} \setminus \Gamma}^1(\mathcal{O})$ so that (5.14) may be used. Then $\vec{v} \cdot \vec{v} \in L^2(\Gamma)$ since $\lambda \in L^2(\Gamma)$ \square

Remark : Note that this lemma simply states that the mixed problem

$$\begin{aligned}
 \Delta \varphi &= r & \text{in } \mathcal{O} \\
 \frac{\partial \varphi}{\partial \vec{n}} &= \lambda & \text{on } \Gamma & \quad \varphi = 0 & \text{on } \partial\mathcal{O} \setminus \Gamma
 \end{aligned}$$

has a unique solution in $H^1(\mathcal{O})$. This is quite normal because if $\phi \in H^1(\mathcal{O})$ it follows that its normal trace is in $H^{-\frac{1}{2}}(\partial\mathcal{O})$.

We can now state, and prove, our main existence result.

Theorem 4.3. *Assume that the permeabilities K_i on the subdomains Ω_i , the permeabilities $K_{i,j}^\tau$ in the fractures $\gamma_{i,j}$ and the thickness $d_{i,j}$ are bounded above and below by a positive constants K_{\min} and K_{\max} , and assume also that $\xi > \frac{1}{2}$. Then the model problem (\mathcal{P}) has a unique solution.*

Proof. To show the existence and uniqueness of the solution, it is sufficient to show that that β satisfies the inf-sup condition, i.e. that there exist a constant C_β such that $\inf_{r \in M} \sup_{v \in V} \frac{\beta(v, r)}{\|r\|_M \|v\|_W} \geq C_\beta$, and to show that α is V-elliptic, i.e. that there exist a constant C_α such that $\alpha(\vec{v}, \vec{v}) \geq C_\alpha \|\vec{v}\|_W^2 \quad \forall \vec{v} \in V$, where V is a subspace defined by

$$V = \{\vec{v} \in W : \beta(\vec{v}, r) = 0 \quad \forall r \in M\},$$

see [13]. Thus the theorem follows from the following two lemmas. \square

Lemma 4.4. *If $\xi > \frac{1}{2}$ then the bilinear form α is V-elliptic.*

Proof. That α is V-elliptic follows from two observations : the first is that if $\vec{v} \in V$ then

$$\|\vec{v}\|_W^2 = \sum_{i \in I} \|\vec{v}_i\|_{0,\Omega_i}^2 + \sum_{(i,j) \in S} \|\vec{v}_{i,j}\|_{0,\gamma_{ij}}^2 + 2 \sum_{(i,j) \in S} \|\vec{v}_i \cdot \vec{\nu}_i + \vec{v}_j \cdot \vec{\nu}_j\|_{0,\gamma_{ij}}^2.$$

This follows from the equations $\beta(\vec{v}, r) = 0 \quad \forall r \in M$, taking successively as test functions

- $r = (\text{div } \vec{v}_1, \text{div } \vec{v}_2, \text{div } \vec{v}_3, 0, 0, 0, 0)$, we obtain $\|\text{div } \vec{v}_i\|_{\Omega_i}^2 = 0$,
- $r = (0, 0, 0, r_{1,2}, r_{2,3}, r_{1,3}, 0)$ with $r_{i,j} = \text{div}_{\gamma_{ij}}(\vec{v}_{i,j}) - (\vec{v}_i \cdot \vec{\nu}_i + \vec{v}_j \cdot \vec{\nu}_j)$, $i \in I, j \in S_i$, we obtain $\|\text{div}_{\gamma_{ij}}(\vec{v}_{i,j})\|_{0,\gamma_{ij}}^2 = \|\vec{v}_i \cdot \vec{\nu}_i + \vec{v}_j \cdot \vec{\nu}_j\|_{0,\gamma_{ij}}^2$,
- $r = (0, 0, 0, 0, 0, 0, \vec{v}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{1,3} \cdot \vec{\nu}_{1,3})$. we obtain $\|\vec{v}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{1,3} \cdot \vec{\nu}_{1,3}\|_{0,T}^2 = 0$.

The second observation is that for each $\vec{v} \in W$

$$\begin{aligned} \alpha(\vec{v}, \vec{v}) &= \sum_{i \in I} \int_{\Omega_i} K_i^{-1} \vec{v}_i \cdot \vec{v}_i + \sum_{(i,j) \in S} \int_{\gamma_{ij}} (d_{i,j} K_{i,j}^\tau)^{-1} \vec{v}_{i,j} \cdot \vec{v}_{i,j} \\ &\quad + \sum_{(i,j) \in S} \int_{\gamma_{i,j}} \frac{d_{i,j}}{2K_{i,j}^n} \left[\xi (\vec{v}_i \cdot \vec{\nu}_i)^2 - 2(1 - \xi) (\vec{v}_j \cdot \vec{\nu}_j) (\vec{v}_i \cdot \vec{\nu}_i) + \xi (\vec{v}_j \cdot \vec{\nu}_j)^2 \right]. \end{aligned}$$

Using this expression we have that for each $\vec{u} \in V$

$$\begin{aligned} \alpha(\vec{u}, \vec{u}) &\geq K_{max}^{-1} \left(\sum_{i \in I} \|\vec{u}_i\|_{0,\Omega_i}^2 + \sum_{(i,j) \in S} \|\vec{u}_{i,j}\|_{0,\gamma_{ij}}^2 \right) \\ &\quad + K_{max}^{-1} \min(1, 2\xi - 1) \sum_{(i,j) \in S} (\|\vec{u}_i \cdot \vec{\nu}_i\|_{0,\gamma_{ij}}^2 + \|\vec{u}_j \cdot \vec{\nu}_j\|_{0,\gamma_{ij}}^2) \\ &\geq K_{max}^{-1} \left(\sum_{i \in I} \|\vec{u}_i\|_{0,\Omega_i}^2 + \sum_{(i,j) \in S} \|\vec{u}_{i,j}\|_{0,\gamma_{ij}}^2 \right) \\ &\quad + K_{max}^{-1} \min(1, 2\xi - 1) \sum_{(i,j) \in S} (\|\vec{u}_i \cdot \vec{\nu}_i + \vec{u}_j \cdot \vec{\nu}_j\|_{0,\gamma_{ij}}^2) \\ &\geq C \|\vec{u}\|_W^2. \end{aligned}$$

where $C = \min(K_{max}^{-1}, \frac{1}{2} K_{max}^{-1} \min(1, 2\xi - 1))$. For the first inequality, we have used the fact that the eigenvalues of the quadratic form $\xi a^2 - 2(1 - \xi)ab + \xi b^2$, where $a, b \in \mathbb{R}$, are 1 and $2\xi - 1$. \square

Lemma 4.5. *The bilinear form β satisfies the inf-sup condition.*

Proof. To show that β satisfies the inf-sup condition, it suffices to show that there is a positive constant C , such that given $r \in M$, there is an element $\vec{v}_r \in W$ such that $\beta(\vec{v}_r, r) = \|r\|_M^2$ and $\|\vec{v}_r\|_W \leq C \|r\|_M$.

Given $r = (r_1, r_2, r_3, r_{1,2}, r_{2,3}, r_{1,3}, r_T) \in M$, to define the first 3 components \vec{v}_1, \vec{v}_2 and \vec{v}_3 of $\vec{v}_r = (\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_{1,2}, \vec{v}_{2,3}, \vec{v}_{1,3})$ let $\tilde{r} \in L^2(\Omega)$ be defined by $\tilde{r}|_{\Omega_i} = r_i$, $i = 1, \dots, 3$, and let $\varphi \in H^2(\Omega)$, be the solution of the auxillary problem

$$\begin{aligned} -\Delta \varphi &= \tilde{r} \quad \text{on } \Omega \\ \varphi &= 0 \quad \text{in } \Gamma. \end{aligned}$$

Put $\vec{v} = -\nabla\varphi$ and, for $i = 1, \dots, 3$, let $\vec{v}_i = \vec{v}|_{\Omega_i}$. Then $\operatorname{div} \vec{v} = \tilde{r}$ so that $\operatorname{div} \vec{v} \in L^2(\Omega)$ and $\vec{v} \in H(\operatorname{div}, \Omega)$. Thus $\vec{v}_i \cdot \vec{\nu}_i|_{\gamma_{ij}} + \vec{v}_j \cdot \vec{\nu}_j|_{\gamma_{ij}} = 0$. Clearly we also have $\vec{v}_i \in H(\operatorname{div}, \Omega_i)$ and $\vec{v}_i \cdot \vec{\nu}_i|_{\gamma_{ij}} + \vec{v}_j \cdot \vec{\nu}_j|_{\gamma_{ij}} \in L^2(\gamma_{ij})$.

To define the next 3 components $\vec{v}_{1,2}$, $\vec{v}_{2,3}$ and $\vec{v}_{1,3}$ of \vec{v}_r , we use Lemma 4.2 and let, for each $(i, j) \in S$, $\varphi_{i,j} \in H_{0,\partial\gamma_{i,j}\setminus T}^1$ be the solution of

$$\int_{\gamma_{i,j}} \nabla\varphi_{i,j} \cdot \nabla\psi \, dx = \int_{\gamma_{i,j}} r_{i,j} \psi \, dx - \int_T \frac{r_T}{3} \psi \, ds \quad \forall \psi \in H_{0,\partial\gamma_{i,j}\setminus T}^1(\gamma_{i,j}). \quad (5.16)$$

Then we let, $\vec{v}_{i,j} = -\nabla_{\gamma}\varphi_{i,j}$. Then $\vec{v} \in H(\operatorname{div}_{\gamma}, \gamma_{i,j})$ with $\operatorname{div} \vec{v}_{i,j} = r_{i,j}$, and $\vec{v}_{i,j} \cdot \vec{\nu} \in L^2(T)$ with $\vec{v}_{i,j} \cdot \vec{\nu} = \frac{r_T}{3}$.

Thus $\vec{v}_r = (\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_{1,2}, \vec{v}_{2,3}, \vec{v}_{1,3}) \in W$, and

$$\begin{aligned} \beta(\vec{v}_r, r) &= \sum_{i \in I} \int_{\Omega_i} \operatorname{div} \vec{v}_i r_i \, dx + \sum_{(i,j) \in S} \left[\int_{\gamma_{ij}} (\operatorname{div}_{\gamma} \vec{v}_{i,j} - (\vec{v}_i \cdot \vec{\nu}_i + \vec{v}_j \cdot \vec{\nu}_j)) r_{i,j} \right] \\ &+ \int_T (\vec{v}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{1,3} \cdot \vec{\nu}_{1,3}) r_T \\ &= \sum_{i \in I} \int_{\Omega_i} r_i^2 + \sum_{(i,j) \in S} \int_{\gamma_{ij}} r_{i,j}^2 + \int_T r_T^2 \\ &= \sum_{i \in I} \|r_i\|_{0,\Omega_i}^2 + \sum_{(i,j) \in S} \|r_{i,j}\|_{0,\gamma_{ij}}^2 + \|r_T\|_{0,T}^2 \\ &= \|r\|_M^2, \end{aligned}$$

and

$$\begin{aligned} \|\vec{v}_r\|_W^2 &= \sum_{i \in I} (\|\vec{v}_i\|_{0,\Omega_i}^2 + \|\operatorname{div} \vec{v}_i\|_{0,\Omega_i}^2) + \sum_{(i,j) \in S} \|(\vec{v}_i \cdot \vec{\nu}_i + \vec{v}_j \cdot \vec{\nu}_j)\|_{0,\gamma_{ij}}^2 \\ &+ \sum_{(i,j) \in S} (\|\vec{v}_{i,j}\|_{0,\gamma_{ij}}^2 + \|\operatorname{div}_{\gamma_{ij}}(\vec{v}_{i,j})\|_{0,\gamma_{ij}}^2) + \|(\vec{v}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{1,3} \cdot \vec{\nu}_{1,3})\|_{0,T}^2 \\ &= \sum_{i \in I} (\|\nabla\varphi_i\|_{0,\Omega_i}^2 + \|r_i\|_{0,\Omega_i}^2) + \sum_{(i,j) \in S} (\|\nabla\varphi_{i,j}\|_{0,\gamma_{ij}}^2 + \|r_{i,j}\|_{0,\gamma_{ij}}^2) + \|r_T\|_{0,T}^2. \end{aligned}$$

However

$$\begin{aligned} \sum_{i \in I} \|\nabla\varphi_i\|_{0,\Omega_i}^2 &= \|\nabla\varphi\|_{0,\Omega}^2 = \int_{\Omega} \nabla\varphi \cdot \nabla\varphi = - \int_{\Omega} \Delta\varphi \cdot \varphi \\ &= \int_{\Omega} r \cdot \varphi \\ &\leq \|r\|_{0,\Omega} \|\varphi\|_{0,\Omega} \\ &\leq \|r\|_{0,\Omega} C_1 \|\nabla\varphi\|_{0,\Omega} \quad (\text{Poincaré}) \end{aligned}$$

and

$$\sum_{i \in I} \|\nabla\varphi_i\|_{0,\Omega_i}^2 = \|\nabla\varphi\|_{0,\Omega}^2 \leq C_1^2 \|r\|_{0,\Omega}^2. \quad (5.17)$$

$$\begin{aligned}
 \|\nabla\varphi_{i,j}\|_{0,\gamma_{ij}}^2 &= \int_{\gamma_{ij}} \nabla\varphi_{i,j} \cdot \nabla\varphi_{i,j} \\
 &= \int_{\gamma_{ij}} r_{i,j} \cdot \varphi_{i,j} + \int_T \frac{r_T}{3} \cdot \varphi_{i,j} \\
 &\leq \|r_{i,j}\|_{0,\gamma_{ij}} \|\varphi_{i,j}\|_{0,\gamma_{ij}} + \|\frac{r_T}{3}\|_{0,T} \|\varphi_{i,j}\|_{0,T}
 \end{aligned}$$

It then follows from the theorem of Poincaré and the trace theorem that there exist positive constants c_2 and c_3 such that

$$\|\nabla\varphi_{ij}\|_{0,\gamma_{ij}}^2 \leq \|r_{i,j}\|_{0,\gamma_{ij}} C_2 \|\nabla\varphi_{ij}\|_{0,\gamma_{ij}} + \|\frac{r_T}{3}\|_{0,T} C_3 \|\nabla\varphi_{ij}\|_{0,\gamma_{ij}}$$

and

$$\|\nabla\varphi_{i,j}\|_{0,\gamma_{ij}} \leq C_2 \|r_{i,j}\|_{0,\gamma_{ij}} + \frac{C_3}{3} \|r_T\|_{0,T}. \quad (5.18)$$

Combining (5.17) and(5.18) we see that for some constant C ,

$$\begin{aligned}
 \|\vec{v}_r\|_W^2 &\leq C \left(\sum_{i \in I} \|r_i\|_{0,\Omega_i}^2 + \sum_{(i,j) \in S} \|r_{i,j}\|_{0,\gamma_{ij}}^2 + \|r_T\|_{0,T}^2 \right) \\
 &\leq C \|r\|_M^2
 \end{aligned}$$

and $\|\vec{v}_r\|_W \|r\|_M \leq C \|r\|_M^2 = C \beta(\vec{v}_r, r)$ so the inf-sup condition is satisfied. \square

5 The discrete formulation of the problem

We assume that the domain Ω is meshed in such a way that the discretization is compatible with the decomposition into subdomains Ω_i and the interfaces are discretized so that the meshes match up at the interface. The intersection T inherits a discretization from the interfaces.

To approximate the problem, the Raviart-Thomas elements of lowest order are used. The approximation spaces $M_{h,i}$, $W_{h,i}$, $M_{h,\gamma_{i,j}}$ and $W_{h,\gamma_{i,j}}$ are subspaces respectively of $L^2(\Omega_i)$, $H(\text{div}, \Omega_i)$, $L^2(\gamma_{i,j})$ and $H(\text{div}_{\gamma_{i,j}}, \gamma_{i,j})$ such that $(M_{h,i}, W_{h,i})$ is a space of lowest order Raviart-Thomas elements on Ω_i , and $(M_{h,\gamma_{i,j}}, W_{h,\gamma_{i,j}})$ is a space of lowest order Raviart-Thomas elements on $\gamma_{i,j}$. The subspace $M_{h,T}$ of $L^2(T)$ is composed of constant functions on each cell of the discretization. The approximation spaces W_h and M_h , finite dimensional subspaces of W and M are defined by

$$\begin{aligned}
 W_h &= (W_{h,1} \times \dots \times W_{h,3} \times W_{h,\gamma_{12}} \times \dots \times W_{h,\gamma_{31}}) \subset W \\
 M_h &= M_{h,1} \times \dots \times M_{h,3} \times M_{h,\gamma_{12}} \times \dots \times M_{h,\gamma_{31}} \times M_{h,T} \subset M.
 \end{aligned}$$

The global discrete problem is written as :

$$\begin{aligned}
 &\text{Find } (\vec{u}_h, p_h) \in W_h \times M_h \text{ such that} \\
 (\mathcal{P}_h) \quad &\alpha(\vec{u}_h, \vec{v}) - \beta(\vec{v}, p_h) = -L_d(\vec{v}), \quad \forall \vec{v} \in W_h \\
 &\beta(\vec{u}_h, r) = L_f(r), \quad \forall r \in M_h.
 \end{aligned}$$

To show the convergence of the discrete problem (\mathcal{P}_h) to the solution (\mathcal{P}) , we need to show that $\alpha : W_h \times W_h \longrightarrow \mathbb{R}$ is uniformly V_h -elliptic, where

$$V_h = \{\vec{v}_h \in W_h : b(\vec{v}_h, r_h) = 0, \quad \forall r_h \in M_h\},$$

and where by uniformly it is meant that the constant of ellipticity for the discrete problems can be chosen independently of h . and that $\beta : W_h \times M_h \longrightarrow \mathbb{R}$ satisfies the uniform discrete inf-sup condition, that is that the constant in the inf-sup condition for the discrete spaces is independent of h . That α satisfies the discrete V_h -ellipticity condition is shown below. However to show the uniform discrete inf-sup condition requires some regularity results that we have not had time to investigate.

To show that $\alpha : W_h \times W_h \longrightarrow \mathbb{R}$ is V_h -elliptic uniformly in h , it suffices to show that $V_h \subset V$. Toward this end, suppose that $\vec{v}_h = (\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_{1,2}, \vec{v}_{2,3}, \vec{v}_{1,3})$ belongs to W_h . From the definition of the approximation spaces it is evident that $\text{div } \vec{v}_{h,i} \in M_{h,i}$, $i=1, \dots, 3$, $\vec{v}_{h,i} \cdot \vec{\nu}_i + \vec{v}_{h,j} \cdot \vec{\nu}_j \in M_{h,\gamma_{i,j}}$ and $\text{div}_\gamma \vec{v}_{h,i,j} \in M_{h,\gamma_{i,j}}$, $(i, j) \in S$, and $\vec{v}_{1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{1,3} \cdot \vec{\nu}_{1,3} \in M_{h,T}$. Thus if $\vec{v}_h \in \vec{W}_h$ we may take as test function in $b(\vec{v}_h, r_h) = 0$, $\forall r_h \in M_h$, the element $r_h = (r_{h,1}, r_{h,2}, r_{h,3}, r_{h,1,2}, r_{h,2,3}, r_{h,1,3}, r_{h,T})$ with $r_{h,i} = \text{div } \vec{v}_{h,i}$, $i = 1, \dots, 3$, $r_{h,i,j} = \text{div}_\gamma \vec{v}_{h,i,j} - (\vec{v}_{h,i} \cdot \vec{\nu}_i + \vec{v}_{h,j} \cdot \vec{\nu}_j)$, $(i, j) \in S$, and $r_{h,T} = \vec{v}_{h,1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{h,2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{h,1,3} \cdot \vec{\nu}_{1,3}$.

One obtains immediately that

$$\text{div } \vec{v}_{h,i} = 0, \quad i = 1, \dots, 3,$$

$$\text{div}_\gamma \vec{v}_{h,i,j} = (\vec{v}_{h,i} \cdot \vec{\nu}_i + \vec{v}_{h,j} \cdot \vec{\nu}_j), \quad (i, j) \in S, \quad \vec{v}_{h,1,2} \cdot \vec{\nu}_{1,2} + \vec{v}_{h,2,3} \cdot \vec{\nu}_{2,3} + \vec{v}_{h,1,3} \cdot \vec{\nu}_{1,3} = 0,$$

and it follows that $\vec{v}_h \in V$ and that $\alpha : W_h \times W_h \longrightarrow \mathbb{R}$ is V_h -elliptic, with an ellipticity constant independent of h .

However to show the uniform discrete inf-sup condition is more delicate. We define a continuous projection operator Π_h from a subspace \mathcal{V} of \vec{W} to \vec{W}_h such that if $\vec{v} \in \mathcal{V}$, $b(\vec{v}, r_h) = b(\Pi_h \vec{v}, r_h)$, $\forall r_h \in M_h$, and to show that for an element r of M , the element $\vec{v}_r \in \vec{W}$ constructed in the previous section in fact belongs to \mathcal{V} , (or to construct another element $\vec{v}'_r \in \mathcal{V}$ having the properties that $b(\vec{v}'_r, r) = \|r\|_M^2$ and $\|\vec{v}'_r\|_{\mathcal{V}} \leq C\|r\|_M$). Here the space \mathcal{V} will be

$$\mathcal{V} = \left\{ v = (\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_{1,2}, \vec{v}_{2,3}, \vec{v}_{1,3}) \in H^1(\Omega_1) \times H^1(\Omega_2) \times H^1(\Omega_3) \right. \\ \left. \times H^1(\gamma_{1,2}) \times H^1(\gamma_{2,3}) \times H^1(\gamma_{1,3}) \right\}.$$

The projection operator $\Pi_h : \mathcal{V} \longrightarrow \vec{W}_h$ is defined by

$$\Pi_h(\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_{1,2}, \vec{v}_{2,3}, \vec{v}_{1,3}) = (\Pi_{h,1}\vec{v}_1, \Pi_{h,2}\vec{v}_2, \Pi_{h,3}\vec{v}_3, \Pi_{h,1,2}\vec{v}_{1,2}, \Pi_{h,2,3}\vec{v}_{2,3}, \Pi_{h,1,3}\vec{v}_{1,3}),$$

where for each i , $i = 1, \dots, 3$, $\Pi_{h,i} : H^1(\Omega_i) \longrightarrow M_{h,i}$ is a Raviart-Thomas projection, and for $(i, j) \in S$, $\Pi_{h,i,j} : H^1(\gamma_{i,j}) \longrightarrow M_{h,i,j}$ is also a Raviart-Thomas

projection. That Π_h has the desired properties follows immediately from the properties of the Raviart-Thomas projection operator. There remains to show that one can construct an element $\vec{v}'_r \in \mathcal{V}$ having the properties that $b(\vec{v}'_r, r) = \|r\|_M^2$ and $\|\vec{v}'_r\|_{\mathcal{V}} \leq C\|r\|_M$.

The first three components of \vec{v}_r as defined in the demonstration of Lemma 4.5 determine a function in $H^1(\Omega)$ if Ω is sufficiently regular and otherwise one can simply solve the problem on a regular domain containing Ω . For the last three components, more regularity results are needed.

6 Domain decomposition for dealing with fractures

In this model, the fractures are represented by interfaces between the various parts of the porous medium and it is natural to use a domaine decomposition method to solve the discretized problem. The unknowns \vec{u}_i , and p_i in the towards that goal, interior of each subdomain are eliminated to bring the problem back to a problem arising on the union of the interfaces. For simplicity, in the remainder of this section, all the problems are written in strong form. Obviously we worked with the weak formulation and in spaces of discretization. Also for simplicity, we suppose the boundary conditions on $\partial\gamma_{i,j} \cap \partial\Omega$ are Dirichlet conditions.

In this part (and in Section 7 for the numerical experiments), we have worked with the model of highly permeable fractures, and thus have replaced equation (5.7) by (5.8) representing the continuity of the pressure along a fracture.

We consider the problem (\mathcal{P}_i) in the subdomain Ω_i

$$\begin{aligned}
 (\mathcal{P}_i) \quad & \operatorname{div}(\vec{u}_i) = f_i && \text{on } \Omega_i \\
 & \vec{u}_i = -K_i \vec{\nabla} p_i && \text{on } \Omega_i \\
 & p_i = p_{di} && \text{in } \Gamma_i \\
 & p_i = p_{i,j} && \text{in } \gamma_{i,j}, \forall j \neq i.
 \end{aligned}$$

By linearity, the solution (\vec{u}_i, p_i) can be written as :

$\vec{u}_i = \vec{u}_i^0 + \vec{u}_i^*$ and $p_i = p_i^0 + p_i^*$ with (\vec{u}_i^0, p_i^0) and (\vec{u}_i^*, p_i^*) the respective solutions of

$$\begin{aligned}
 (\mathcal{P}_i^0) \quad & \operatorname{div}(\vec{u}_i^0) = 0 && \text{on } \Omega_i \\
 & \vec{u}_i^0 = -K_i \vec{\nabla} p_i^0 && \text{on } \Omega_i \\
 & p_i^0 = 0 && \text{in } \Gamma_i \\
 & p_i^0 = p_{i,j} && \text{in } \gamma_{i,j}. \\
 (\mathcal{P}_i^*) \quad & \operatorname{div}(\vec{u}_i^*) = f_i && \text{on } \Omega_i \\
 & \vec{u}_i^* = -K_i \vec{\nabla} p_i^* && \text{on } \Omega_i \\
 & p_i^* = p_{di} && \text{in } \Gamma_i \\
 & p_i^* = 0 && \text{in } \gamma_{i,j}.
 \end{aligned}$$

The equations (5.9), \dots , (5.12) on the interface $\gamma_{i,j}$ can be written as

$$\begin{aligned}
 (\mathcal{P}_{\gamma_{i,j}}) \quad & \operatorname{div}_{\gamma_{i,j}} \vec{u}_{i,j} = f_{i,j} + (\vec{u}_i^* \cdot \vec{\nu}_i + \vec{u}_j^* \cdot \vec{\nu}_j) + (\vec{u}_i^0 \cdot \vec{\nu}_i + \vec{u}_j^0 \cdot \vec{\nu}_j) && \text{in } \gamma_{i,j} \\
 & \vec{u}_{i,j} = -d_{i,j} K_{i,j}^\tau \vec{\nabla}_{\gamma_{i,j}} p_{i,j} && \text{in } \gamma_{i,j} \\
 & p_{i,j} = p_{d\gamma_{i,j}} && \text{in } \partial\gamma_{i,j} \cap \partial\Omega \\
 & p_{i,j} = p_T && \text{in } T.
 \end{aligned}$$

By eliminating the unknown $\vec{u}_{i,j}$, the first two equations of this system can be replaced by the equation

$$\operatorname{div}_{\gamma_{i,j}}(-d_{i,j}, K_{i,j}^\tau \vec{\nabla}_{\gamma_{i,j}} p_{i,j}) = f_{i,j} + (\vec{u}_i^* \cdot \vec{\nu}_i + \vec{u}_j^* \cdot \vec{\nu}_j) + (\vec{u}_i^0 \cdot \vec{\nu}_i + \vec{u}_j^0 \cdot \vec{\nu}_j) \quad \text{in } \gamma_{i,j}.$$

To write the problem as a problem posed on the interface $\gamma = \bigcup_{(i,j) \in S} \gamma_{i,j}$. We define the space

$$M_f = L^2(\gamma) \times L^2(T) = L^2(\gamma_{12}) \times L^2(\gamma_{23}) \times L^2(\gamma_{31}) \times L^2(T).$$

Then $\lambda \in M_f$ is written as : $\lambda = (\lambda_\gamma, \lambda_T) = (\lambda_{1,2}, \lambda_{2,3}, \lambda_{1,3}, \lambda_T)$ we write γ_i for $\partial\Omega_i \cap \gamma = \bigcup_{j \in S_i} \gamma_{i,j}$.

For each subdomain Ω_i , the Steklov-Poincaré type operator \mathcal{S}_i is defined so that for $\lambda \in M_f$,

$$\mathcal{S}_i(\lambda) = \mathcal{S}_i(\lambda_i) = -(\vec{u}_i^0 \cdot \vec{\nu}_i)|_{\gamma_i},$$

where (\vec{u}_i^0, p_i^0) is the solution of problem (\mathcal{P}_i^0) which $p_{i,j} = \lambda_{i,j}, \forall j \in S_i$, and the operator χ_i

$$\chi_i(f_i, p_{di}) = (\vec{u}_i^* \cdot \vec{\nu}_i)|_{\gamma_i},$$

where (\vec{u}_i^*, p_i^*) is the solution of problem (\mathcal{P}_i^*) .

By extending $\mathcal{S}_i(\lambda)$ and χ_i to all of γ with an extension by 0 on $\gamma \setminus \gamma_i$, we can define

$$\mathcal{S}(\lambda) = \sum_{i \in I} \mathcal{S}_i(\lambda) \quad \text{and} \quad \chi(f, p_d) = \sum_{i \in I} \chi_i(f_i, p_{di})$$

and write the global system on the global interface γ

$$(\mathcal{P}_\gamma) \quad \begin{aligned} \mathcal{S}(\lambda) - \operatorname{div}_\gamma(d_\gamma K_\gamma^\tau \vec{\nabla}_\gamma \lambda) &= f_\gamma + \chi(f, p_d) && \text{in } \gamma \\ \lambda &= p_{d_\gamma} && \text{on } \partial\gamma \cap \partial\Omega \\ \lambda_{i,j} - \lambda_T &= 0 && \text{on } T, i \in I, j \in S_i \\ \mathcal{S}_T(\lambda) &= 0 && \text{in } T, \end{aligned}$$

where d_γ, K_γ^τ and f_γ are defined in an obvious way, and the operator \mathcal{S}_T is defined by

$$\mathcal{S}_T(\lambda) = \sum_{i \in I, j \in S_i} (-d_{i,j} K_{i,j}^\tau \vec{\nabla}_\gamma \lambda_{i,j})|_T \cdot \nu_{i,j},$$

so that the last equation of the global system represents the continuity of the flow imposed on T in equation (5.13).

This system is symmetric and positive definite. Thus a simple conjugate gradient itération can be used to solve the linear system. Each iteration requires the solution of a Dirichlet problem in each subdomain, and a calculation of the flow in the fracture.

7 Numerical results

7.1 Description the test cases

In this part we will present some numerical results to illustrate the properties of our model. In the numerical experiments, we consider a domain Ω that contains in addition to the 3 fractures of the model problem (Figure 5.2), a simple isolated fracture. See (Figure 5.3).

Three test cases are considered. In each of the three cases, a pressure equal to 1 imposed on top of the fractures and a pressure equal to 0 on the bottom of the fractures. In all the domains, the permeability is supposed to be equal to 1, whereas on the fractures the permeability is supposed to be much greater than 1 as the model used for experiments is the model for more permeable fractures (with pressure continuous across the fractures).

The test case : horizontal flow

We present an example of the results with and without fractures. We impose a horizontal pressure gradient : $p = 1$ on the left lateral boundary and $p = 0$ on the right lateral boundary and a no flow condition on the other external boundaries of the subdomains. For boundary conditions on the fractures, we impose the Dirichlet condition $p = 1$ on the upper part of the fracture and $p = 0$ on the lower part of the boundaries and a no flow condition on the rest of the fracture boundaries.

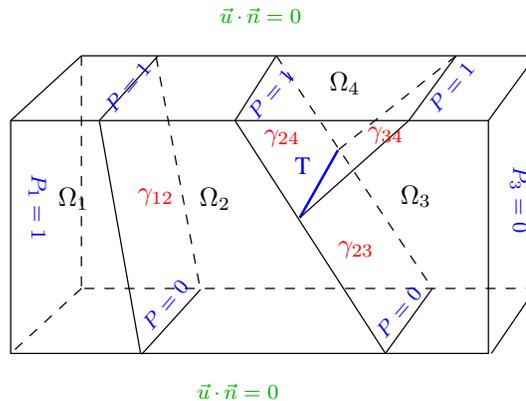


FIG. 5.3: Figure representing the geometry of the first case test

The permeability in the fracture multiplied by its thickness is equal to 100. This means that the fluid tends to flow rapidly along the fracture.

Figure (5.4) represents the pressure given in the absence of the fractures (we represented each subdomain separately). The flow is only horizontal, the pressure varies linearly from left to right. If we take into account the fractures, the situation changes, as is shown in Figure (5.5). We see in addition to the horizontal flow, a visible effect of the flow in the fractures.

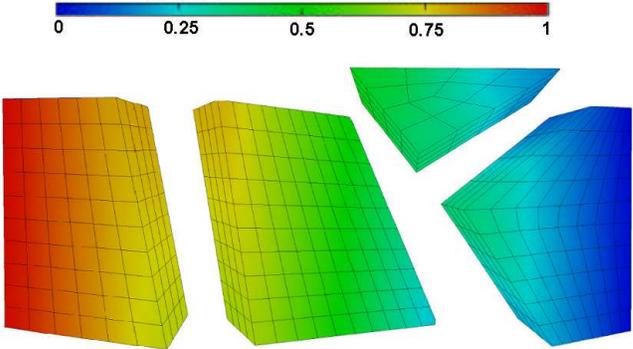


FIG. 5.4: The pressure given by the model without fractures

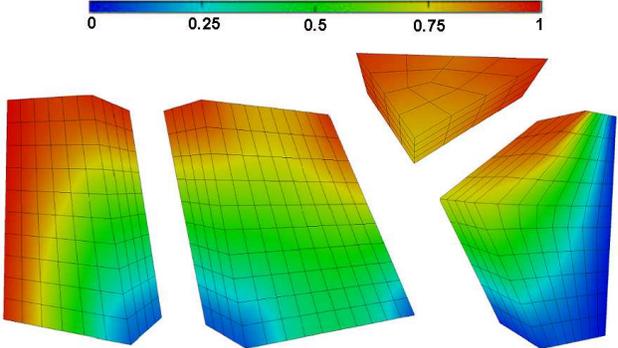


FIG. 5.5: The pressure given by the model with fractures

Results : Flow due to the fractures

In the second test case, the only modification in the data is that a pressure equal to 1 is imposed on both sides of Ω . Without fractures the pressure would be constant (the figure is not included), but with fractures Figure (5.6) shows a significant effect due to the fractures.

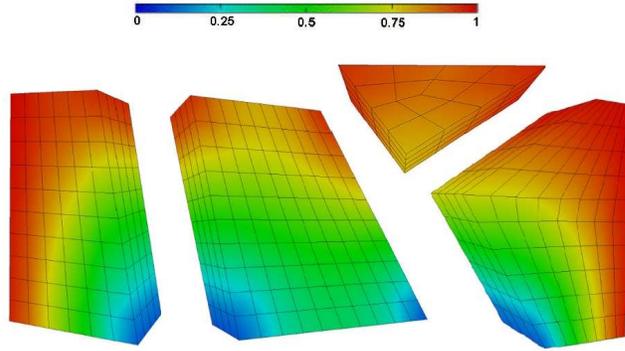


FIG. 5.6: The pressure given by the model with fractures

By comparing Figure (5.5) and Figure (5.6), one also sees the difference due to the boundary condition imposed on the right.

Results : Vertical flow

For the third test we impose a vertical pressure gradient. Figure (5.7) shows the

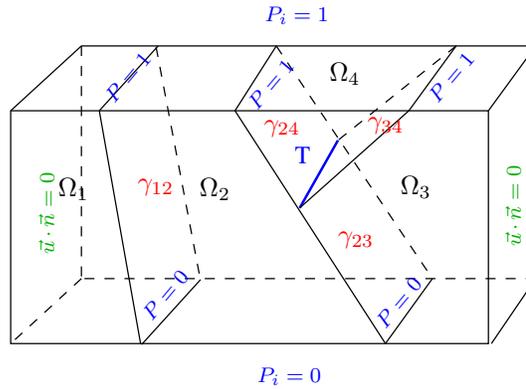


FIG. 5.7: Figure representing the geometry of the case third test

geometry and the boundary conditions for this test case : $p = 1$ on the top boundary, $p = 0$ on the bottom boundary and a no flow condition on the other boundaries.

Figure (5.8) represents the pressure field in the absence of fractures. We see a linear pressure drop from the top to the bottom, from 1 to 0.

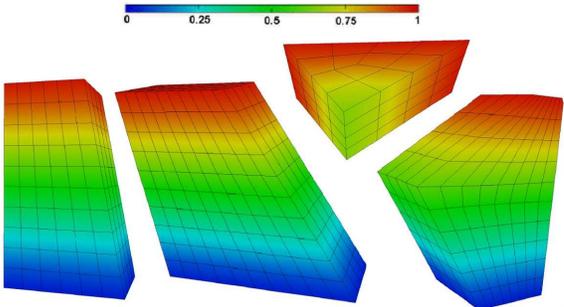


FIG. 5.8: The pressure given by the model without fractures

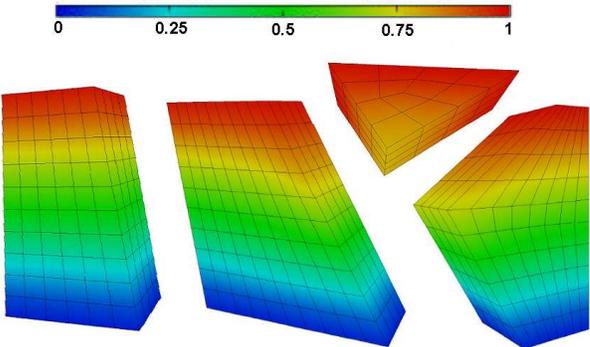


FIG. 5.9: The pressure given by the model with fractures

By comparing these results with those given in the presence of the fractures Figure (5.9), we see a difference on the interfaces, in the presence of fractures, one notices in addition to the vertical flow in the subdomains a contribution coming from the fractures.

7.2 Performance of the preconditioner

In this paragraph we implement a new preconditioner to accelerate the convergence of the conjugate gradient method used to solve the problem.

In the problems without fractures, for the classical method of domain decomposition, the Neumann–Neumann preconditioner has been shown to be efficient [85]. However, if we compare the differential orders of the different operators of the system in the global interface problem (\mathcal{P}_γ) , it's natural to take as preconditioner the inverse of the flow operator in the fracture $(\operatorname{div}_\gamma(d_\gamma K_\gamma^\tau \vec{\nabla}))^{-1}$. It is of order 2, while the Steklov–Poincaré operator is only of order 1.

Depending on the value of the dK_γ^τ , one of the other preconditioners, or a combination of the two could possibly be effective.

In the case of domain decomposition with fractures, we applied a conjugate gradient with and without the preconditioner, the inverse of the flow operator in the fracture. The result is shown in Figure (5.10) that represents the residual as a function of the number of iterations. Without preconditioner, the convergence is very slow, and by using a preconditioner, we notice a dramatic improvement in the convergence. The number of iterations in the convergence is divided by 10 when using the preconditioner.

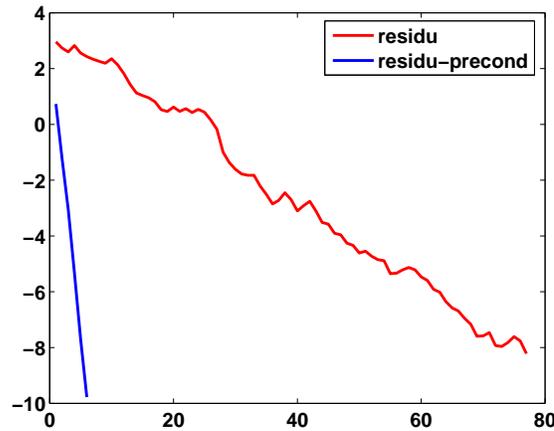


FIG. 5.10: Efficiency of the preconditioner

For the second test case, we applied the conjugate gradient method with preconditioner, for various values of dK_γ .

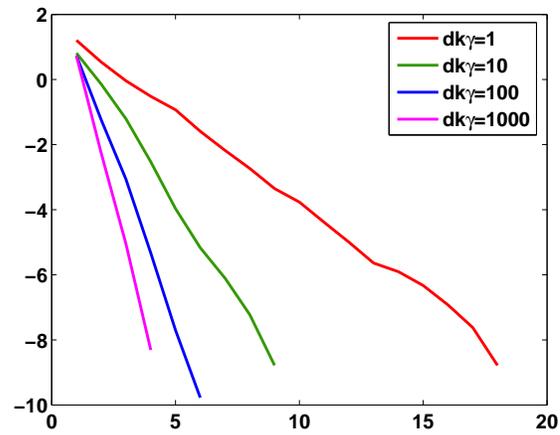


FIG. 5.11: Efficiency of the preconditioner depending on dK_γ

The results in Figure (5.11) show that the preconditioner works better if the value of dK_γ is higher.

8 Conclusion

In this paper we have studied a model for flow in a porous medium with fractures, in which the fractures are treated as interfaces. In particular we have been concerned with the case of intersecting fractures. Existence and uniqueness of the solution in this case has been shown. A discretization method using mixed finite elements has been proposed. The model was reformulated as a domain decomposition problem. The method was implemented for a 3D problem and numerical results were shown. A preconditioner was introduced and numerical results showed its usefulness.

Conclusion générale

Au cours de cette thèse, nous avons étudié une nouvelle méthode de couplage global entre le transport et la chimie, basée sur une méthode de Newton-Krylov.

Pour réaliser cette étude, une première partie de cette thèse a été consacré à l'étude et la résolution du problème d'advection-diffusion avec différents schémas et méthodes numériques. La résolution de l'équation d'advection-diffusion en milieux poreux nécessite la connaissance préalable du champ de l'écoulement. L'utilisation de la méthode EFMH pour le calcul du champ d'écoulement assure une description de l'écoulement beaucoup plus précise que les méthodes d'EF de Lagrange. La technique de séparation d'opérateurs est utilisée pour discrétiser l'équation d'advection-diffusion. La partie diffusive est approchée en utilisant la méthode EFMH et la partie advective est approchée par la méthode VF. L'intérêt de cette technique est de traiter des processus différents comme l'advection, la diffusion, la chimie, etc. avec des méthodes numériques appropriées, et ainsi d'économiser du temps de calcul en prenant des échelles de temps différentes pour la diffusion et l'advection.

La deuxième partie a été consacrée à la modélisation du problème de la chimie sous l'hypothèse d'équilibre chimique. On a introduit les notions de composants et d'espèces secondaires, qui permettent de formuler un système chimique à l'équilibre.

Pour coupler le transport et la chimie, la majorité des méthodes proposées se base sur un couplage pas à pas : on résout itérativement la chimie et le transport. Ce procédé converge, mais parfois au prix d'une réduction draconienne du pas de temps. Le système couplé transport-chimie est décrit par un système d'équations non linéaires. La difficulté est bien évidemment la taille du système sous-jacent, dont le nombre d'inconnues est le nombre de points de grille multiplié par le nombre d'espèces chimiques. Par ailleurs, une mise en oeuvre naïve de cette méthode (en formant le jacobien du système) est impossible pour des systèmes réalistes, et ne permet pas de conserver les codes de chimie et de transport séparés. Il en résulte que les temps de calculs de ce type de simulation s'élèvent rapidement avec la complexité du problème simulé jusqu'à devenir prohibitifs. Pour remédier à ce problème, on utilise une méthode de Newton-Krylov pour résoudre le système non-linéaire. Il s'agit d'une méthode de Newton dans laquelle les systèmes linéaires sont résolus par une méthode itérative (plus souvent GMRES ou BiCGSTAB). Ces méthodes, aussi appelées "Jacobian free" demandent seulement le produit Jacobien-vecteur sans former et stocker la matrice Jacobienne.

Finalement, une partie de cette thèse a été consacré à l'étude d'un modèle d'écoulement pour les fractures qui s'intersectent dans les milieux poreux. Les fractures sont assimilées à des interfaces entre les sous domaines par une méthode de décomposition de domaines. Nous avons démontrés l'existence et l'unicité de la solution. Des résultats numériques ont été présentées pour la validation du modèle étudié, d'autres résultats numériques ont été présentés pour valider et montrer l'efficacité du preconditionneur choisi.

Perspectives

Les résultats préliminaires que nous avons présentés sont prometteurs : il est possible de résoudre efficacement des problèmes géochimiques en utilisant une méthode de Newton-Krylov, bien qu'il reste plusieurs points qui doivent être abordés.

- le premier est d'appliquer la méthode sur des cas tests de configurations plus complexes, en 2D ou 3D, où nous nous attendons à ce que la méthode montre sa grande capacité.
- il sera alors certainement nécessaire d'explorer la question de choisir et d'utiliser un préconditionneur pour réduire le nombre de Krylov itérations. Vu la structure en bloc de la matrice Jacobienne, il est naturel de réutiliser les méthodes de séparations d'opérateurs, comme celles proposées par [45].
- les résultats présentés précédemment sont donnés pour un pas de temps fixe, qui était insuffisant pour une intégration sur grand intervalle. Pour résoudre avec succès les problèmes difficiles comme le benchmark ci-dessus, il est clairement nécessaire d'utiliser des pas de temps adaptatifs.

Pour poursuivre le travail de thèse sur la modélisation des fractures, il serait intéressant de valider le modèle d'écoulement pour les fractures qui s'intersectent sur un test réel et aussi d'étudier le modèle du transport.

Publications

Journal

- L. Amir, M. Kern, V. Martin, et J. Roberts ,
'Décomposition de domaine pour un milieu poreux fracturé : un modèle en 3D avec fractures qui s'intersectent',
ARIMA, Vol.5, pp. 11-25. Nov 2006.
- L. Amir, M. Kern,
'Newton-Krylov methods for coupling transport with chemistry in heterogeneous porous media', soumis dans le journal Computational Geosciences 2008.
- L. Amir, M. Kern et J. E. Roberts,
'Theoretical and numerical study for a model of intersecting fractures in a porous medium', to submit.

Conférences, Séminaires et Proceedings

- 'Coupling transport with chemistry by using Newton-Krylov methods',
CMWR XVII - International Conference on Computational Methods in Water Resources XVII, San Francisco, USA, July 6-10 2008.
- 'Reactive transport in heterogeneous porous media',
Slovak-Austrian Mathematical Congress, Podbanske, Slovakia, September 16-21, 2007.
- 'Newton-Krylov methods for fully coupling chemistry and transport',
6th ICIAM International congress on Industrial and Applied mathematics, Zurich, Switzerland, July 16-20, 2007.
- 'A Globally Coupled Method for Reactive Transport in Porous Media',
SIAM Conference on Mathematical and Computational Issues in the Geosciences, Santa Fe, New Mexico , USA, March 19-22, 2007.
- 'Newton-Krylov methods for coupling transport with chemistry in porous media',
CMWR XVI - International Conference on Computational Methods in Water Resources XVI, Copenhagen, Denmark, (editors : P. Binning, P. Engesgaard, H. Dahle, G. Pinder, W. G. Gray), June 2006
- 'Les méthodes de Newton-Krylov pour le couplage transport chimie',
Congrès CANUM (38ème Congrès National d'Analyse Numérique), à Beg er Lenn Guidel Morbihan, France, du 29 mai au 2 juin 2006.
- 'Décomposition de domaine et préconditionnement pour un modèle 3D en milieu poreux fracturé',
8èmes Journées d'Analyse Numérique et Optimisation, Rabat (Maroc), December 14-16, 2005.
- 'Domain decomposition for a fractured porous medium',
8th SIAM Conference on Mathematical and Computational Issues in the Geosciences, Avignon (France), June 7-10, 2005.

- 'Décomposition de domaine pour un milieu poreux fracturé',
2ème colloque sur les Tendances dans les Applications Mathématiques en Tunisie, Algérie, Maroc, Tunis, April 26-28, 2005.
- 'Reactive transport in porous media',
Séminaire organisé à l'occasion d'un séjour à l'Université d'Erlangen (Allemagne, groupe du Prof. P. Knabner, 25 au 29 juin 2007), dans le cadre d'un partenariat Hubert Curien Procope.
- 'Full coupling in transport of solutes in porous media with chemistry',
Workshop3+3, Numerical computing for groundwater flows, Mohammadia school engineering, Rabat, Morocco, september 20-22 2006.

Bibliographie

- [1] C. Alboin, J. Jaffré, and J. E. Roberts. Domain decomposition for flow in fractured porous media. *Domain Decomposition Methods in Sciences and Engineering, Domain Decomposition Press*, pages pp. 365–373.
- [2] C. Alboin, J. Jaffré, J. E. Roberts, and C. Serres. Modeling fractures as interfaces for flow and transport in porous media. *Contemporary Mathematics*, pages Volume 295, pp. 13–24, 2002.
- [3] L. Amir and M. Kern. Newton-krylov methods for coupling transport with chemistry in porous media. In P. Binning, P. Engesgaard, H. Dahle, G. Pinder, and W. G. Gray, editors, *XVI International Conference on Computational Methods in Water resources (CMWR XVI)*, 2006.
- [4] L. Amir and M. Kern. Newton-Krylov methods for coupling transport with chemistry in heterogeneous porous media. *submitted in Computational Geosciences*, 2008.
- [5] L. Amir, M. Kern, V. Martin, and J. E. Roberts. Décomposition de domaine pour un milieu poreux fracturé : un modèle en 3d avec fractures qui s’intersectent. *ARIMA*, pages Vol.5, pp. 11–25, November.
- [6] L. Amir, M. Kern, and J. E. Roberts. Theoretical and numerical study for a model of intersecting fractures in a porous medium. to submit.
- [7] M. Bercovier, O. Pironneau, and V. Sastri. Finite elements and characteristics for some parabolic-hyperbolic problems. *Appl. Math. Modelling*, 7,89 :96, 1983.
- [8] C. Bernardi, Y. Maday, and A.T. Patera. Domain decomposition by mortar element method. *Technical report, Publications du laboratoire d’analyse numérique de l’université de Pierre et Marie Curie, R92013*,.
- [9] Craig M. Bethke. *Geochemical Reaction Modelling*. Oxford University Press, 1996.
- [10] D. Billaux, B. Paris, and C. Darcel. *3FLO Version 2.1. Calculs d’écoulements et de transport tridimensionnels. Rapport Interne ITASCA Consultants, Volumes 1 à 4*. 2004.
- [11] N. Bouillard. *Développement de méthodes numérique pour le transport réactif*. PhD thesis, Université de Provence, 2006.

- [12] Pierre Brassard and Paul Bodurtha. A feasible set for chemical speciation problems. *Computers in Geosciences*, 26 :277–291, 2000.
- [13] F. Brezzi. *On the existence, uniqueness and approximation of saddle point problems arising from Lagrangian multipliers*, *PAIRO, Anal, Numér*, 8. 1974.
- [14] F. Brezzi, Jim Douglas, Jr., M. Fortin, and D. Marini. Efficient rectangular mixed finite elements for second order elliptic problems in two and three variables. *M2AN*, 51, 581 :604, 1987.
- [15] F. Brezzi, Jim Douglas, Jr., and D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, 47 :217-235, 1985.
- [16] F. Brezzi and M. Fortin. Mixed and hybrid finite element method. In *Springer-Verlag, New York*, 1991.
- [17] J. Carrayrou. *Modélisation du transport de solutés réactifs en milieu poreux saturé*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, 2001.
- [18] J. Carrayrou. GDR MoMaS benchmark – reactive transport. published electronically at http://www.gdrmomas.org/ex_qualifications.html, 2006.
- [19] J. Carrayrou. GDR MoMaS benchmark – reactive transport. *Computational Geosciences*, 2008. submitted.
- [20] G. Chavent and J. Jaffré. Mathematical models and finite elements for reservoir simulation. 1986.
- [21] G. Chavent and J. E. Roberts. A unified physical presentation of mixed, mixed-hybrid finite element method and standard finite difference approximations for the determination of velocities in waterflow problems. *Adv. Water Resour.*, 14(6), 329 :348, 1991.
- [22] R. Dautray and J.L. Lions. *Analyse mathématique et calcul numérique*, volume 4. Masson, Paris, 1988.
- [23] C. Dawson and M. F. Wheeler. Time splitting methods for advection-diffusion-reaction equations arising in contaminant transport. *Industrial and Applied Mathematics, Siam*, 1992.
- [24] C. de Dieuleveult and J. Erhel. A numerical model for coupling chemistry and transport. In *International Conference on SCientific Computation And Differential Equations, SciCADE 2007*, 2007.
- [25] C. de Dieuleveult, J. Erhel, and M. Kern. A global strategy for solving reactive transport equations. In preparation, 2008.
- [26] J-R. de Dreuzy, P. Davy, J. Erhel, and J. Bremond d’Ars. Anomalous diffusion exponents in continuous two-dimensional multifractal media. / *Physical Review E* /, pages Volume 70, 016306 (6 pages), 2004.
- [27] J-R. de Dreuzy and J. Erhel. Efficient algorithms for the determination of the connected fracture network and the solution to the steady-state flow equation in fracture networks. *Computers and Geosciences*, pages Volume 29, pp. 107–111, 2003.

-
- [28] G. de Marsily. *Hydrogéologie quantitative, collection des Sciences de la Terre*. Masson, Paris, 1981.
- [29] Jim Douglas, Jr. and T. Arbogast. Dual porosity models for flow in naturally fractured reservoirs. *Dynamics of Fluids in Hierarchical Porous Formations*, pages J. H. Cushman (ed.), Academic Press, pp. 177–221, 1990.
- [30] Jim Douglas, Jr., T. Arbogast, and U. Hornung. Derivation of the double porosity model of single phase flow via homogenization theory. *SIAM J. Math. Anal.*, pages 21 823–836, 1990.
- [31] Jim Douglas, Jr. and J. E. Roberts. Global estimates for mixed methods for second order elliptic equations. *Math. Comp.*, 44, 39 :52, 1985.
- [32] Jim Douglas, Jr. and T. F. Russell. Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM J. Numer. Anal.* 19, 871 :885, 1982.
- [33] D. A. Dzombak and F. M. M. Morel. *Surface Complexation Modelling, Hydrous Ferric Oxides*. Wiley Interscience, 1990.
- [34] D.A. Dzombak and F.M.M. Morel. *Surface complexation modeling—Hydrous ferric oxide*. John Wiley, 1993.
- [35] S. C. Eisenstat and H. F. Walker. Choosing the forcing terms in an inexact Newton method. *SIAM Journal on Scientific Computing*, 17(1) :16–32, 1996.
- [36] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In P. G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis*, volume VII, pages 713–1020. North-Holland, 2000.
- [37] M. Fahs, A. Younes, J. Carrayrou, and Ph. Ackerer. Efficiency of the direct substitution approach for multicomponent reactive transport in porous media. *Water, Air, & Soil Pollution*, pages 299–308, 2008.
- [38] Leij Feike and J. H. Dane. Analytical solution of the one dimensional advection equation equation and two- or tree-dimentional dispersion equation. *Water Ressources Res*, 26, 1475-1482, 1990.
- [39] Roland W. Freund and Noël M. Nachtigal. QMR : a quasi-minimal residual method for non-Hermitian linear systems. *Numerische Mathematik*, 60 :315–339, 1991.
- [40] N. Frih, J. E. Roberts, and A. Saada. Un modèle darcy-forchheimer pour un écoulement dans un milieu poreux fracturé. *ARIMA*, pages Vol.5, pp. 129–143, November.
- [41] V. Girault and P. A. Raviart. *Finite element approximation of the Navier-Stokes equations*. In : *Lectures Notes on Mathematics*, Springer Verlag, volume 749.
- [42] Gilles Gnacadja. Fixed points of order-reversing maps in $\mathbb{R}_{>0}^n$ and chemical equilibrium. *Math. Methods Appl. Sci.*, 30(2) :201–211, 2007.

- [43] C. P. L. Grady, Jr. and H. C. Lin. *Biological Wastewater Treatment : Theory and Applications*. Marcel-Dekker, 1980.
- [44] G. E. Hammond. *Innovative methods for solving multicomponent biogeochemical groundwater transport on supercomputers*. PhD thesis, University of Illinois at Urbana-Champaign, 2003.
- [45] G. E. Hammond, A. J. Valocchi, and P. C. Lichtner. Application of Jacobian-free Newton–Krylov with physics-based preconditioning to biogeochemical transport. *Advances in Water Resources*, 28 :359–376, 2005.
- [46] J. Hoffmann, S. Kräutle, and P. Knabner. A parallel global implicit 2-d solver for reactive transport problems in porous media based on a reduction scheme and its application to the MoMaS benchmark problem. *Computational Geosciences*, 2009. article soumis.
- [47] H. Hoteit. *Simulation d’écoulements et de transports de polluants en milieu poreux : Application à la modélisation de la sûreté des dépôts de déchets radioactifs*. PhD thesis, Université Rennes 1, 2002.
- [48] H. Hoteit, Ph. Ackerer, and R. Mose. Nuclear waste disposal simulations : Couplex test cases : Simulation of transport around a nuclear waste disposal site : The COUPLEX test cases (editors : Alain Bourgeat and Michel Kern). *Computational Geosciences*, 8 :99–124, 2004.
- [49] R. Mosé J. Carrayrou and Ph. Behra. Operator-splitting procedures for reactive transport and comparison of mass balance errors. *Journal of Contaminant Hydrology*, 68(3-4) :239–268, 2004.
- [50] J. F. Kanney, C. T. Miller, and C. T. Kelley. Convergence of iterative split operator approaches for approximating nonlinear reactive transport problems. *Advances in Water Resources*, 26(247-261), 2003.
- [51] K. Karlsen and K. Risebo. An operator splitting method for convection diffusion equations. *Numer. Math*, 1997.
- [52] C. T. Kelley. *Iterative methods for linear and nonlinear equations*, volume 16 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. With separately available software.
- [53] C. T. Kelley. *Solving nonlinear equations with Newton’s method*. Fundamentals of Algorithms. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003.
- [54] D. A. Knoll and D. E. Keyes. Jacobian-free newton-Krylov methods : a survey of approaches and applications. *J. Comput. Phys.*, 193(2) :357–397, 2004.
- [55] S. Kraeutle and P. Knabner. A new numerical reduction scheme for fully coupled multicomponent transport-reaction problems in porous media. *Water Resources Research*, 41(W09414), 2005.
- [56] S. Kraeutle and P. Knabner. A reduction scheme for coupled multicomponent transport-reaction problems in porous media : Generalization to problems with

- heterogeneous equilibrium reactions. *Water Resources Research*, 43(W03429), 2007.
- [57] V. Lagneau and J. van der Lee. HYTEC results of the momas reactive transport benchmark. *Computational Geosciences*, 2008. submitted.
- [58] J.L. Lions and E. Magenes. *Non-homogeneous Boundary Value Problems and Applications*, volume 1. Springer, Berlin, 1972.
- [59] J. David Logan. *Transport Modeling in Hydrogeochemical Systems*. Springer-Verlag, New York, 2001.
- [60] P.-L. Lucille, A. Burnol, and Ph. Ollar. Chemtrap : a hydrogeochemical model for reactive transport in porous media. *Hydrological processes*, 14 :2261–2277, 2000.
- [61] V. Martin, J. Jaffré, and J. E. Roberts. Modeling fractures and barriers as interfaces for flow in porous media. *SIAM Journal on Scientific Computing*, pages vol. 26, No 5, pp. 1667–1691 (electronic), 2005.
- [62] U. Meissner. A mixed finite element model for use in potential flow problem. *Int. J. Numer. Methods Eng.*, 6, 467 :473, 1973.
- [63] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.
- [64] F. M. M. Morel and J. G. Hering. *Principles and Applications of Aquatic Chemistry*. Wiley, New-York, 1993.
- [65] F.M.M. Morel and J.G. Hering. *Principles and Applications of Aquatic Chemistry*. Wiley, New-York, 2nd edition, 1993.
- [66] J.-C. Nedelec. A new family for mixed finite elements r3. *Numer. Math.* 35, 315 :341, 1980.
- [67] J.-C. Nedelec. Mixed finite elements in r3. *Numer. Math.*, 50,57 :82, 1986.
- [68] David L. Parkhurst and C.A.J. Appelo. User’s guide to PHREEQC (version 2)- A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. Technical Report 99-4259, USGS, 1999.
- [69] J.-P. Sauty R. Fabriol and G. Ouzounian. Coupling geochemistry with a particle tracking transport model. *Journal of Contaminant Hydrology*, 13, 117-129, 1993.
- [70] P.-A. Raviart and J.-M. Thomas. A mixed hybrid finite element method for the second order elliptic problem. *Lectures Notes in Mathematics 606*, Springer-Verlag, New York, 292 :315, 1977.
- [71] J. E. Roberts and J.-M. Thomas. Mixed and hybrid methods. In P. Ciarlet and J.-L. Lions, editors, *Handbook of numerical analysis II, Finite element methods 1*, 1989.
- [72] Jacob Rubin. Transport of reacting solutes in porous media : Relation between mathematical nature of problem formulaion and chemical nature of reaction. *Water Ressources Research*, 19(5) :1231–1252, October 1983.

- [73] Y. Saad and M. H. Schultz. GMRES : a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3) :856–869, 1986.
- [74] M. Saaltink, C. Ayora, and J. Carrera. A mathematical formulation for reactive transport that eliminates mineral concentrations. *Water Resources Research*, 34(7) :1649–1656, 1998.
- [75] M. Saaltink, J. Carrera, and C. Ayora. On the behavior of approaches to simulate reactive transport. *Journal of Contaminant Hydrology*, 48 :213–235, 2001.
- [76] M.W. Saaltink, J. Carrera, and C. Ayora. A comparison of two approaches for reactive transport modelling. *Journal of Geochemical Exploration*, 69-70 :97–101, 2000.
- [77] A.-L. Salignac. *Transport multi-espèces et réactions géochimiques en aquifère : développement et validation du modèle couplé HYTEC 2D*. PhD thesis, École des Mines de Paris, 1998.
- [78] A. Sboui. *Quelques méthodes numériques robustes pour l'écoulement et le transport en milieu poreux*. PhD thesis, Université Paris Dauphine, 2007.
- [79] N. Z. Shapiro and L. S. Shapley. Mass action laws and the gibbs free energy function. *J. SOC. Indust. Appl. Math.*, 13(2) :353–375, 1965.
- [80] P. Siegel, R. Mosé, Ph. Ackerer, and J. Jaffré. Solution of the advection-dispersion equation using a combination of discontinuous and mixed finite elements. *Journal for Numerical Methods in Fluids*, 24 :595–613, 1997.
- [81] Laura Sigg, Philippe Behra, and Werner Stumm. *Chimie des milieux aquatiques*. Dunod, 3^{ème} edition, 2000.
- [82] W. R. Smith and R. W. Missen. *Chemical Reaction Equilibrium Analysis : Theory and Algorithms*. Wiley Interscience, 1982. reprinted with corrections, Krieger, Malabar, FL., 1991.
- [83] G. Strang. On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, 5, pages 506–517, 2006.
- [84] Ne-Zheng Sun. *Mathematical modeling of groundwater pollution*. Springer-Verlag and Geological Publishing House, New York, 1996.
- [85] P. Le Tallec. Domain decomposition methods in computational mechanics. *Computational Mechanics Advances*, pages vol. 1, No 2, pp. 122–220.
- [86] C. V. Theis. *The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage*, volume 16. Am. Geophys. Union Trans., 1935.
- [87] J.-M. Thomas. *Sur l'Analyse Numérique des Méthodes d'Elément Finis Hybrides et Mixtes*. PhD thesis, Thèse de Doctorat d'Etat, Université de Pierre et Marie Curie, 1977.

-
- [88] E. Tillier. *Couplage réactions transport pour la modélisation et la simulation du stockage géologique de co₂*. PhD thesis, Université de Marne-La-Vallée, 2007.
- [89] L. Trenty, A. Michel, Y. Le Gallo, and E. Tillier. A sequential splitting strategy for co₂ storage modeling. In *10th European Conference on the Mathematics of Oil Recovery, Amsterdam*, 2006.
- [90] J. van der Lee. CHESS, another speciation and surface complexation computer code. Technical Report LHM/RD/93/39, CIG École des Mines de Paris, Fontainebleau, 1993.
- [91] Jan van der Lee. Thermodynamic and mathematical concepts of CHESS. Technical Report LHM/RD/98/39, CIG – Ecole des Mines de Paris, Fontainebleau, France, 1998.
- [92] H. A. van der Vorst. BI-CGSTAB : A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems. 13(2) :631–644, March 1992.
- [93] G. T. Yeh and V. S. Tripathi. A critical evaluation of recent developments in hydrogeochemical transport models of reactive multichemical components. *Water Resources Research*, 25 :93–108, 1989.
- [94] A. Younes. Modélisation du transport réactif en milieux poreux non saturés avec une méthode ellam en maillage variable. *C. R. Acad. Sci., Géosciences*, 336, pages 547–552, 2004.
- [95] A. Younes. A moving grid eulerian lagrangian localized adjoint method for solving one-dimensional nonlinear advection-diffusion-reaction equations. *Transp. Porous Media*, pages 60, 241–250, 2005.

