HAL open science

Analyse mathématique de modèles géochimiques

Tangi Migot, Jocelyne Erhel

To cite this version:

Tangi Migot, Jocelyne Erhel. Analyse mathématique de modèles géochimiques. [Rapport de recherche] INRIA Rennes, équipe SAGE. 2014. hal-01107638

HAL Id: hal-01107638

https://hal.science/hal-01107638v1

Submitted on 21 Jan 2015

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.

Analyse mathématique de modèles géochimiques

Tangi Migot and Jocelyne Erhel
équipe SAGE, INRIA

Analyse mathématique de modèles géochimiques

Résumé : Dans ce document est étudié la modélisation et la résolution d'un système géochimique. On utilise une modélisation espèces/composants dont découle des lois d'action de masse et des lois de conservation de la masse. On s'intéresse principalement aux phénomènes de précipitation-dissolution qui sont représentés dans le modèle par un problème de complémentarité. Sont ensuite discutées deux approches : la première utilise l'aspect polynomial du système et essaye de trouver des solutions "analytique" avec une méthode algébrique, et l'autre approche numérique qui utilise une fonction de reformulation pour utiliser des méthodes de Newton semi-lisse.

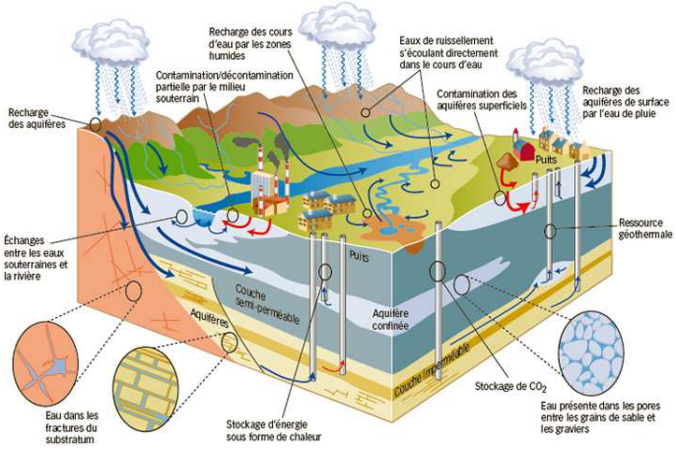
Mots-clés : Newton semi-smooth - Précipitation-Dissolution - Système polynomial - Géochimie - Problème de complémentarité

Introduction

Contexte

L'eau est un élément indispensable pour tous les organismes vivants sur la Terre. Plus de 70% de la surface de la Terre est recouverte d'eau, ceci essentiellement sous forme liquide stockée dans différents réservoirs principalement les océans, les glaciers et les eaux souterraines. Néanmoins cette eau essentielle à la vie est menacée par la pollution, les rejets industriels ou l'utilisation nocive de pesticides dans l'agriculture.

FIGURE 1 – https://interstices.info/ressources-naturelles Mieux gérer les ressources naturelles à Vincent Landrin (d'après Natural Resources Canada).



Cette ressource étant nécessaire à la vie il est normal de vouloir la préserver et c'est une des préoccupations principale des sciences de la terre. Ces dernières décennies les scientifiques de la terre et les modélisateurs ont collaboré pour créer des modèles physiques afin de réaliser des simulations et comprendre les différents risques liés à l'eau. Parmi ces risques on peut mentionner le drainage minier acide.

Le drainage minier acide est une solution acide qui apparaît dans les eaux à la suite de l'exploitation d'une mine ou du stockage de déchets miniers. Lorsque l'on exploite une mine, des minéraux sont exposés à l'eau et à l'air ce qui provoque des réactions lentes qui à terme créent de l'acide sulfurique qui va contaminer les aquifères. En effet rien que l'acidification de l'eau est mortelle pour bien des êtres vivants, quasiment pas de vie en dessous d'un ph de 5. De plus les métaux qui réagissent avec l'eau peuvent devenir des poisons violents tel que l'arsenic. Une des caractéristiques de la pollution par le drainage minier acide est le dépôt de rouille dans l'eau qui donne une couleur rouge.

FIGURE 2 – Broken Banks, Bishop Auckland, Co Durham, UK Younger (1995) Q. Jl. Eng. Geol. 28: S101 - S113

Ces dernières décennies de grand projets scientifiques ont été lancé pour lutter contre ce type de pollution. L'objectif étant de pouvoir prédire et identifier la formation du drainage minier acide ainsi que tester de nouvelles méthodes pour le neutraliser.

Motivations

Le transport réactif en milieu poreux est l'un de ces phénomènes qui ont attiré l'attention des scientifiques. Il s'agit du déplacement d'espèces chimiques qui peuvent réagir entre elles dans un milieu poreux. Le phénomène de transport est accompagné d'advection et de diffusion. Le transport réactif a déjà été utilisé dans le contexte du drainage minier acide [19] et au sein de l'équipe SAGE ce problème a été étudié dans le cadre de la gestion de déchets radioactifs [6, 5, 30]. La modélisation du transport réactif en milieu poreux est un problème difficile car elle représente un couplage entre le phénomène physique de transport et le phénomène de chimie, chacun possédant ses propres difficultés. Cette modélisation conduit généralement à un système d'équations aux dérivées partielles et algébriques. Plusieurs méthodes existent pour résoudre ce problème, par exemple : méthode séquentielle non itérative (SNIA), méthode séquentielle itérative (SIA), méthode d'approche globale (GDAE) qui avait été choisi par [30].

Le travail de ce stage a été concentré autour du phénomène chimique. En effet, il est parfois nécessaire d'analyser plus finement une partie du problème pour en déduire des propriétés numériques utiles pour la réalisation de simulations. L'objectif est donc de discuter la modélisation du système de chimie avec un intérêt particulier pour la prise en compte des réactions de précipitation-

dissolution qui amènent un problème de complémentarité ainsi que d'analyser différentes approches de résolution de ce phénomène : d'une part une approche analytique qui utilise l'aspect polynomial du système et d'autre part une approche numérique qui utilise une méthode de Newton semi-lisse.

Plan

Ce rapport est organisé en 3 parties.

La première partie est consacrée à la modélisation du système de chimie. Dans un premier temps on reprend les notions essentielles sur la modélisation espèces/composants et les différentes hypothèses que l'on a posé. Ensuite on a choisi de modéliser les phénomènes de sorption et de précipitation-dissolution. Pour obtenir en fin le système de chimie complet avec l'hypothèse que les réactions sont à l'équilibre. Par la suite dans ce rapport on ne considérera pas les réactions de sorption.

En seconde partie, on montre 2 exemples pour lesquels on utilise l'aspect polynomial du système afin d'en obtenir les solutions en analysant un à un les différents états du système, i.e en fonction de la présence ou non des minéraux. Une fois ces 2 exemples analysés on proposera une généralisation de cette méthode suivi d'une application à un exemple à 6 minéraux utilisé dans la littérature.

Enfin la troisième partie est concentrée sur la résolution numérique du système de chimie. Pour ce faire on rappellera l'utilisation de la méthode de Newton semi-lisse, rappel qui sera précédé d'une étude sur le jacobien et notamment la question de son inversibilité. Par la suite nous montrerons des exemples de modifications de la méthode de Newton semi-lisse pour prendre en compte la positivité de certaines variables.

Enfin nous verrons comment l'approche analytique peut nous aider à développer le calcul numérique et nous conclurons sur ce travail et les nouvelles perspectives.

Contents

I Modèle de Chimie **8**

1 Introduction **8**

2 Réaction à l'équilibre **9**

3 Cinétique chimique **10**

4 Modélisation espèces et composants **10**

5 Lois de conservation **11**
5.1 Conservation de la masse 11
5.2 Conservation de charge 11

6 Modèle chimique **11**
6.1 Hypothèses 11
6.2 Système chimique 12
6.2.1 Réactions aqueuses et surfaciques 12
6.2.2 Réactions de précipitation-dissolution 12
6.2.3 Lois de conservation de la masse 13
6.2.4 Système chimique complet 13

II Étude semi-analytique **14**

7 Chimie avec 2 sels : Chlorure de Sodium et Chlorure de Potassium **14**
7.1 Variables 15
7.2 Mise en équation 15
7.3 Cas limite : toutes les inégalités sont actives 16
7.4 Solutions 16
7.5 Graphique 19

8 Chimie avec 3 sels : Chlorure de Sodium, Chlorure de Potassium et Carnallite **19**
8.1 Variables 20
8.2 Mise en équation 20
8.3 Cas limite : toutes les inégalités sont actives 21
8.4 Solutions 22
8.5 Graphique 22

9 Généralisation de l'approche analytique **22**
9.1 Point critique 23
9.2 Courbes critiques 23
9.3 Résolution du système polynomial 24
9.4 Application numérique 25

10 Chimie avec 6 sels : évolution de DMA en présence de 6 minéraux **26**
10.1 Matrice stœchiométrique 28
10.2 Mise en équation 28
10.3 Cas limite : toutes les inégalités sont actives 29
10.4 Solutions 29

11 Conclusion **31**

12 Ajout des réactions de sorption **32**

III Étude numérique **33**

13 Étude du jacobien généralisé **33**

14 Méthode de Newton semi-lisse **35**

15 Prise en compte des contraintes **37**
15.1 Kanzow Active-Set 37
15.2 Modification du pas 41
15.3 Newton avec pénalisation de la fonction objectif 42
15.4 Comparaison des méthodes 43

16 Choix du point initial **45**

IV Conclusion **48**

V Appendice **50**

A Calculs des solutions : exemple 3 sels **50**

B Benchmark MoMas Transport Réactif "Easy Test Case" **55**

Références **57**

Part I
Modèle de Chimie

Dans cette partie est décrit le modèle mathématique utilisé pour représenter un système de chimie. Ce type de formulation est notamment utilisé dans le cadre du transport réactif, par exemple dans les thèses [14] et [30].

1 Introduction

On considère N_E espèces chimiques pour N_r réactions qui donnent des réactions chimiques exprimées sous la forme :

$$\sum_{t=1}^{N_e} \nu_{r,t} E_t \rightleftharpoons 0 \ , r = 1, ..., N_r$$

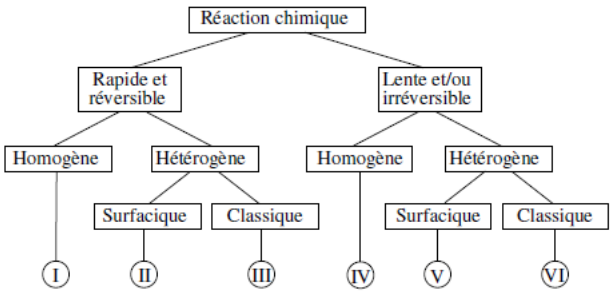
avec $\nu_{r,t}$ les coefficients de la matrice stœchiométrique. On suppose ces réactions chimiques indépendantes ; dans le cas contraire il est possible de réduire le système de façon à ce que les réactions deviennent indépendantes.

On peut considérer l'évolution de ces réactions chimiques de deux façons différentes [27, 8, 3]:

- La réaction chimique est rapide et on atteint l'équilibre local assez vite. La réaction est réversible. On peut ainsi supposer l'équilibre présent à chaque instant. On obtient des équations algébriques non-linéaires, plus facile à résoudre et avec moins d'inconnues dans le système.

- La réaction chimique est lente et dans ce cas on doit tenir compte de l'évolution temporelle de la réaction. Il faut ici utiliser une approche cinétique par des équations différentielles ordinaires. La réaction est irréversible.

[27] propose une classification plus détaillée à l'intérieur de chaque catégorie :



L'approche utilisée dépend du critère utilisé pour déterminer si une réaction est "rapide" ou non. Dans le contexte du transport réactif, une réaction est considérée lente si sa vitesse de réaction est du même ordre de grandeur ou

plus lente que la vitesse du phénomène physique de transport. L'approche la plus couramment utilisée est une approche mixte : une partie des équations considérées à l'équilibre et les autres font intervenir la cinétique [28, 18, 16, 29, 20, 25, 10].

Dans la pratique, deux modèles chimiques sont utilisés pour représenter ces réactions chimiques :

- Lois d'action de masse, qui seront utilisées dans la suite de ce chapitre.
- La minimisation contrainte de l'enthalpie libre (aussi appelée énergie de Gibbs).

La différence étant principalement le type de problème qui en découle ; dans le premier on a un système non-linéaire et dans le second un problème d'optimisation sous contrainte. Ces deux modèles sont similaires conceptuellement et numériquement, voir [21].

2 Réaction à l'équilibre

Sont considérées dans cette section des réactions chimiques supposées être dans un équilibre local. Lorsque les réactions chimiques sont supposées à l'équilibre, **les lois d'action de masse** donnent :

$$K_r - \prod_{i=1}^{N_E} a_i^{\nu_{r,i}} = 0,\ r = 1, N_r$$

où a_i est l'activité de l'espèce E_i et K_r la constante d'équilibre de la réaction r.

Activité d'une espèce aqueuse : L'activité d'une espèce aqueuse E_i est fonction de sa molalité m_i qui exprime la quantité de l'espèce E_i en solution (moles/kg de solvant) : $a_i = \gamma_i * m_i / m_i^0$, où γ_i est le coefficient d'activité et m_i^0 est la molalité de référence (égale à 1 mol/kg d'eau). Pour des concentrations faibles le coefficient d'activité peut être supposé égal à 1 et donc si le solvant est l'eau on a : $a_i = m_i$. Cette hypothèse est raisonnable pour la modélisation du système chimique. Pour des milieux plus concentrés, il faut approximer le coefficient d'activité, par exemple loi de Debye-Hückel.

Activité de l'eau : Si la solution est diluée, l'activité de l'eau est voisine de 1 et n'intervient donc pas dans les lois d'action de masse, sinon voir loi de Raoult.

Activité d'une espèce fixée : pas de modèle utilisé uniformément. Si la solution est idéale on peut utiliser comme représentation de l'activité la fraction molaire [9] ou également supposer que l'activité est égale à la molarité (nombre de moles d'une espèce par litre) [13]. Dans le cas d'une solution non-idéale il faut tenir compte d'un coefficient d'activité [7].

Activité d'une espèce précipitée : pour un minéral l'activité est toujours égale à 1.

3 Cinétique chimique

Sont considérées dans cette section les réactions chimiques où l'approche cinétique est utilisée. En fonction du type de réaction (homogène, hétérogène) on utilise des lois de vitesse pour calculer la vitesse de réaction R_k de la réaction k, ce qui donne des équations différentielles. Le modèle utilisé est détaillé dans [26]. D'autres modèles peuvent être utilisés pour modéliser des réactions cinétiques comme dans le cas de phénomènes de biodégradation avec le modèle de Monod [9] et ses variantes.

Dans le cadre de phénomènes de précipitation-dissolution, ce sont des réactions hétérogènes qui sont concernées. Ces lois cinétiques nous permettent dans un premier temps de calculer le taux d'avancement de la réaction (donné en mol.s^{-1} ou en mol.year^{-1}) pour un précipité p_i, ν_{p_i}:

$$v_{p_i} = K_+ \prod a_j^{E_{ij}} \left[1 - \left(\frac{Q}{K}\right)^m\right]$$

- K_+: constante cinétique de la réaction
- a_j: activité de l'espèce j
- Q: produit des activités ioniques : $Q = \frac{\prod_j [c_j]^{s_j}}{\prod_j [\bar{c_j}]^{r_j}}$ les $\bar{c}$ sont les composants réactifs et les c les composants produits
- K: produit de solubilité du minéral (constante qui dépend de la température)
- m: souvent égal à 1

On remarque qu'à l'équilibre $Q = K$ la vitesse de réaction est nulle, la vitesse est positive si la solution est sursaturée $Q > K$ et négative si la solution est sous-saturée $Q < K$. Ensuite on utilise ce taux pour obtenir des équations différentielles. Soit $\nu \in \mathbb{R}^n$ le vecteur de n taux d'avancement pour n minéraux. Pour chaque espèce (aqueuse ou minérale) e, on a :

$$\frac{de_i}{dt} = \sum_j^n E_{ij}\nu_j$$

En allant de t' à un nouveau point t, le taux instantané va évolué au fur et à mesure que le fluide chimique évolue. Plutôt que d'utiliser le taux au temps t' sur toute l'étape, il est plus précis de prendre la moyenne des taux au temps t' et t.

4 Modélisation espèces et composants

L'hypothèse d'un système chimique avec des réactions indépendantes et les lois d'action de masse permet d'exprimer une partie des inconnues (espèces secondaires) du système en fonction d'autres inconnues (composants). De même

que pour la loi d'action de masse, pour chaque équation cinétique on associe une espèce cinétique qui n'apparaitra que dans la réaction cinétique et les réactions à l'équilibre. Ensuite l'activité de chaque espèce permet de répartir espèces aqueuses, espèces fixées et espèces précipitées. L'activité des espèces précipitées ne dépendant pas du nombre de moles du précipité, on ne peut pas l'exprimer en fonction d'autres inconnues. Cette répartition espèces/composants est très utilisée par les modélisateurs géochimiques, voir [3, 10]. Les inconnues du système sont les concentrations des composants aqueux (c), fixés (s) et précipités (p). Sont notés α et β les concentrations des espèces secondaires aqueuses et fixées et on note ($cine$) la concentration des espèces cinétiques.

5 Lois de conservation

5.1 Conservation de la masse

Les équations de conservation de la masse expriment la conservation de la masse en terme de composants et d'espèces cinétiques. Elles consistent à définir les masses totales en composants et en espèces cinétiques comme la répartition de la masse de l'ensemble des espèces du système. Les masses totales sont invariantes par les équations d'équilibre chimique mais dépendant de la vitesse de réaction des réactions cinétiques.

5.2 Conservation de charge

Une solution aqueuse est toujours électriquement neutre, elle possède donc autant de charges positives que négatives. Ce principe de neutralité électrique s'écrit : $\sum_i z_i m_i = 0$

6 Modèle chimique

Le modèle de chimie décrit ici est utilisé dans l'équipe SAGE dans le cadre du transport réactif développé par [30], [14] mais aussi [6] et[5].

6.1 Hypothèses

On considère un milieu saturé en eau, on suppose que les espèces en solution sont suffisamment diluées pour que la solution soit supposée idéale. Les espèces considérées sont les espèces aqueuses et solides, aucune espèce gazeuse et cinétique. Les phénomènes pris en compte sont les réactions aqueuses, les réactions de sorption-désorption qui correspondent aux réactions de surface et les réactions de précipitation-dissolution. Pour les espèces aqueuses et fixées, l'activité est donc directement remplacée par la concentration molaire (molalité), par hypothèse que la densité du fluide est constante. De plus on suppose que les réactions sont à l'équilibre chimique.

6.2 Système chimique

6.2.1 Réactions aqueuses et surfaciques

Les lois d'action de masse permettent d'obtenir le système d'équations suivant

$$\begin{cases} \alpha_i = K_\alpha \prod_{l=1}^{N_c} c_l^{S_{i,l}} & i = 1, ..N_\alpha \\ \beta_j = K_\beta \prod_{l=1}^{N_c} c_l^{A_{j,l}} \prod_{m=1}^{N_s} s_m^{B_{j,m}} & j = 1, ..N_\beta \end{cases}$$

avec le tableau de Morel [15] suivant $\begin{bmatrix} & c \in \mathbb{R}^{N_c} & s \in \mathbb{R}^{N_s} \\ \alpha \in \mathbb{R}^{N_\alpha} & S & 0 \\ \beta \in \mathbb{R}^{N_\beta} & A & B \end{bmatrix}$.

Plus souvent utilisé sous une forme logarithmique :

$$\begin{cases} log(\alpha_i) = K_\alpha + \sum_{l=1}^{N_c} S_{ij} log(c_l) & = 1, ..N_\alpha \\ log(\beta_j) = K_\beta + \sum_{l=1}^{N_c} A_{jl} log(c_l) + \sum_{m=1}^{N_s} B_{jm} log(s_m) & j = 1, ..N_\beta \end{cases}.$$

Les inconnues deviennent $x = log(c)$ et $y = log(s)$, l'avantage est de garantir la positivité stricte des variables.

Par exemple présente dans [17] mais aussi dans [14] et [30] ou [4]. Cette forme logarithmique est utilisée dans plusieurs logiciels comme SPECY, MIN3P, RETRASO, RICHY, parce que cette forme assure la positivité des concentrations. De plus, cette forme facilite le calcul des dérivées car elle linéarise les équations : au lieu de calculer la dérivée des produits et des puissances, on calcule la dérivée des sommes.

Dans [30] on a vu quelques défauts de cette formulation : singularité de la jacobienne dans le cas d'un traceur ou de concentrations faibles et meilleures performances numériques sans les logarithmes. De plus la formulation sans logarithme nous donne un système polynomial qui nous permet d'autres approches.

6.2.2 Réactions de précipitation-dissolution

Les réactions de précipitation-dissolution sont des réactions hétérogènes. Les espèces aqueuses ont la possibilité à un seuil de saturation de précipiter et ainsi de former un minéral ou inversement un minéral peut se dissoudre en espèces aqueuses. Il est difficile de dire a priori si un minéral est présent ou non dans la solution, il faut donc considérer 2 cas :

- le fluide est saturé et il y a précipitation

$$\begin{cases} 0 = (1 - K_i \prod_{j=1}^{N_c} c_l^{E_{i,j}}) \\ 0 < p_i \end{cases} \quad i = 1 \dots N_p$$

- le fluide est sous-saturé et le minéral est entièrement dissous

$$\begin{cases} 0 \leq (1 - K_i \prod_{j=1}^{N_c} c_l^{E_{i,j}}) \\ 0 = p_i \end{cases} \quad i = 1 \dots N_p$$

où on suppose qu'il n'y a qu'un seul minéral par réaction, K_i est le produit de solubilité de la i-ème réaction et la matrice $E \in \mathbb{R}^{N_p \times N_c}$ contient les coefficients des réactions de précipitation-dissolution.

Une méthode récente dans le contexte du transport réactif [4], [24] mais bien connue en optimisation est d'exprimer ce phénomène par un problème de complémentarité

$$(PC)\ \varphi(a,b) = 0 \iff ab = 0 \text{ et } a \geq 0\ b \geq 0.$$

Pour résoudre ce type de problème une astuce est d'utiliser une fonction de reformulation, par exemple en utilisant la fonction $\varphi_{min}(a,b) = min(a,b)$.

$$(PC) \iff min(a,b) = 0$$

La formule qui regroupe les deux cas possibles est la suivante :

$$\varphi(\Pi_i(c), p_i) = \begin{cases} \Pi_i(c) = K_i - \prod_{j=1}^{N_c} c_l^{E_{i,j}} & \text{si } \Pi_i(c) \leq p_i \\ p_i & \text{si } \Pi_i(c) > p_i \end{cases} i = 1 \ldots N_p$$

6.2.3 Lois de conservation de la masse

On exprime les concentrations totales pour chaque composant, total aqueux T_c et total fixé T_s sont exprimés tel que :

$$\begin{cases} T_c(c,s,p) & = c + S^t\alpha(c) + A^t\beta(c,s) + E^t p \\ T_s(c,s) & = s + B^t\beta(c,s) \end{cases}$$

Les lois de conservation de masse sont alors écrites sous la forme :

$$\begin{cases} T & = T_c(c,s,p) \\ W & = T_s(c,s) \end{cases}$$

où T et W sont des données dans le cas d'un système fermé.

6.2.4 Système chimique complet

Au final, on obtient le système chimique suivant :

$$\begin{cases} T - T_c(c,s,p) & = 0 \\ W - T_s(c,s) & = 0 \\ \alpha_i = K_\alpha \prod_{l=1}^{N_c} c_l^{S_{i,l}} & i = 1,..N_\alpha \\ \beta_j = K_\beta \prod_{l=1}^{N_c} c_l^{A_{j,l}} \prod_{m=1}^{N_s} s_m^{B_{j,m}} & j = 1,..N_\beta \\ \varphi(\Pi(c), p) & = 0 \end{cases}$$

qui est un système fermé de $(N_c + N_s + N_p)$ inconnues et autant d'équations, avec des contraintes de positivité sur les inconnues c, s et p pour conserver le sens physique du système.

Dans la suite de ce rapport nous allons considérer le modèle chimique où apparaissent uniquement les réactions de précipitation-dissolution. Dans la dernière section est faite un lien entre l'étude du modèle précipitation-dissolution et le modèle général qui vient d'être vu.

Modèle utilisé On va considérer dans cette partie des problèmes qui ne concernent que les phénomènes de précipitation et dissolution où est faite l'hypothèse qu'un minéral n'intervient que dans une seule réaction. Les coefficients de ces réactions sont regroupés dans la matrice stœchiométrique E :

$$\begin{bmatrix} & concentrations: \ c \\ précipités: p & E \end{bmatrix}$$

On applique la loi d'action de masse (minérale) qui donne un problème de complémentarité :

$$min(K_i - \prod_{j=1}^{N_c} c_j^{E_{ij}}, p_i) = 0, \ i = 1 \dots N_p$$

où N_p est le nombre de précipités et N_c le nombre d'espèces.

De plus, on a les lois de conservations qui forment un système de N_c équations.

$$T - c - E^t p = 0$$

Le système de chimie est un système fermé, il y a le même nombre d'équations $(N_c + N_p)$ que d'inconnues c et p.

On peut ajouter la loi de conservation de charges qui nous donne une relation entre les données T, c'est la "neutralité électrique". On a ainsi $(N_c - 1)$ données.

Part II
Étude semi-analytique

Ont été étudiés 3 exemples de système chimique : deux exemples avec 2 sels et 3 sels qui sont des exemples classiques de réactions chimiques qui vont nous permettre de généraliser une démarche de résolution analytique que l'on appliquera dans le dernier exemple de 6 sels qui est un exemple provenant de la littérature dans le contexte du drainage minier acide.

7 Chimie avec 2 sels : Chlorure de Sodium et Chlorure de Potassium

On s'intéresse à la réaction de 2 précipités dans une solution : le Chlorure de Sodium et le Chlorure de Potassium.

$$\begin{aligned} Na^+ + Cl^- &\rightleftharpoons NaCl \\ K^+ + Cl^- &\rightleftharpoons KCl \end{aligned}$$

7.1 Variables

Soient c_1, c_2 et c_3 les variables de concentrations de Na^+, Cl^- et K^+ et p_1, p_2 les variables de concentrations des précipités $Nacl$ et Kcl, de plus K_1 et K_2 sont les produits de solubilités spécifiques aux réactions (constantes). Par interprétation physique des variables du problème les c et p sont des variables positives.

7.2 Mise en équation

Lois d'action de masse minérale

$$\begin{aligned} min(K_1 - c_1c_3, p_1) &= 0 \\ min(K_2 - c_2c_3, p_2) &= 0 \end{aligned}$$

Lois de conservation de masse

$$\begin{aligned} T_1 &= c_1 + p_1 \\ T_2 &= c_2 + p_2 \\ T_3 &= c_3 + p_1 + p_2 \end{aligned}$$

Neutralité électrique

$$T_3 = T_1 + T_2$$

On élimine les variables c_1, c_2 et c_3 pour obtenir :

$$\begin{aligned} min(K_1 - (T_1 - p_1)(T_1 + T_2 - p_1 - p_2), p_1) &= 0 \\ min(K_2 - (T_2 - p_2)(T_1 + T_2 - p_1 - p_2), p_2) &= 0 \end{aligned}$$

On peut aussi l'écrire comme un système avec 4 inégalités (où il y a toujours 2 = et et 2 $\leq$):

$$\begin{cases} K_1 - (T_1 - p_1)(T_1 + T_2 - p_1 - p_2) & \geq 0 \\ K_2 - (T_2 - p_2)(T_1 + T_2 - p_1 - p_2) & \geq 0 \\ p_1 & \geq 0 \\ p_2 & \geq 0 \end{cases} \tag{1}$$

Pour rappel, T_1 et T_2 sont des paramètres, K_1 et K_2 des constantes et les inconnues sont p_1 et p_2.

On aurait également pu choisir d'exprimer notre système en fonction des variables c_1 et c_2, dans ce cas on considèrerait le système suivant :

$$\begin{aligned} min(K_1 - c_1(c_1 + c_2), T_1 - c_1) &= 0 \\ min(K_2 - c_2(c_1 + c_2), T_2 - c_2) &= 0 \end{aligned}$$

7.3 Cas limite : toutes les inégalités sont actives

On cherche s'il existe une solution au problème où toutes les inégalités du système sont des égalités, on a alors :

$$\begin{cases} K_1 - (T_1 - p_1)(T_1 + T_2 - p_1 - p_2) & = 0 \\ K_2 - (T_2 - p_2)(T_1 + T_2 - p_1 - p_2) & = 0 \\ p_1 & = 0 \\ p_2 & = 0 \end{cases}$$

Il existe une unique solution à ce système qui vérifie les contraintes de positivité pour p :

$p_1 = 0$, $p_2 = 0$ et $T_1^p = \frac{K_1}{\sqrt{K_1+K_2}}$, $T_2^p = \frac{K_2}{\sqrt{K_1+K_2}}$

7.4 Solutions

Il y a $2^{N_p} = 2^2 = 4$ façons de choisir une combinaison de 2 égalités et 2 inégalités dans le système (1), chacune des ces combinaisons correspond à une situation dans lequel se trouve le système, par exemple les 2 précipités sont présents ou ils sont dissous ou un seul des deux est présent, etc. On étudie chacun des cas et à chaque fois les 2 égalités nous donnent les solutions du problème c_1 et c_2 en fonction des constantes et des paramètres. Ensuite, les 2 inégalités nous donnent des contraintes sur les valeurs de T_1 et T_2 qu'il faut respecter pour être dans ce cas.

Cas 1 : 2 sels $p_1 > 0$ et $p_2 > 0$ Le système devient :

$$\begin{cases} K_1 - (T_1 - p_1)(T_1 + T_2 - p_1 - p_2) & = 0 \\ K_2 - (T_2 - p_2)(T_1 + T_2 - p_1 - p_2) & = 0 \\ p_1 & > 0 \\ p_2 & > 0 \end{cases}$$

En utilisant les expressions en fonction de c_1 et c_2 on obtient :

$$\begin{cases} c_1(c_1+c_2) & = K_1 \\ c_2(c_1+c_2) & = K_2 \\ c_1 & < T_1 \\ c_2 & < T_2 \end{cases}$$

Les solutions sont $c_1 = \frac{K_1}{\sqrt{K_1+K_2}} = T_1^p$, $c_2 = \frac{K_2}{\sqrt{K_1+K_2}} = T_2^p$ et donc $p_1 = T_1 - T_1^p$, $p_2 = T_2 - T_2^p$ sous les contraintes :

$$\begin{cases} T_1 & > T_1^p \\ T_2 & > T_2^p \end{cases}$$

Cas 2 :0 sel $p_1 = 0$ **et** $p_2 = 0$ Le système devient :

$$\begin{cases} K_1 - (T_1 - p_1)(T_1 + T_2 - p_1 - p_2) & > 0 \\ K_2 - (T_2 - p_2)(T_1 + T_2 - p_1 - p_2) & > 0 \\ p_1 & = 0 \\ p_2 & = 0 \end{cases}$$

Les solutions sont $p_1 = 0$ et $p_2 = 0$, à condition de vérifier les inégalités :

$$\begin{cases} T_1(T_1+T_2) < & K_1 \\ T_2(T_1+T_2) < & K_2 \end{cases}$$

Remarque 1 : On peut aussi écrire $T_1(T_1+T_2) < K_1$ et $T_2(T_1+T_2) < K_2$ sous la forme $x^2 + xT_i - K_j < 0$ les racines sont : $x_1 = \frac{1}{2}(-T_i + \sqrt{T_i^2 + 4K_j}) > 0$ et $x_2 = \frac{1}{2}(-T_i - \sqrt{T_i^2 + 4K_j}) < 0$. Le polynôme est négatif entre les racines.

Cas 3 :1 sel (p1) $p_1 > 0$ **et** $p_2 = 0$ Le système devient :

$$\begin{cases} K_1 - (T_1 - p_1)(T_1 + T_2 - p_1 - p_2) & = 0 \\ K_2 - (T_2 - p_2)(T_1 + T_2 - p_1 - p_2) & > 0 \\ p_1 & > 0 \\ p_2 & = 0 \end{cases}$$

En utilisant les expressions en fonction de c_1 et c_2 on obtient :

$$\begin{cases} c_1(c_1+c_2) & = K_1 \\ c_2(c_1+c_2) & < K_2 \\ c_1 & < T_1 \\ c_2 & = T_2 \end{cases}$$

On a le polynôme de degré 2 considéré dans la Remarque 1 en c_1: $c_1(c_1 + T_2) - K_1 = 0$ qui d'après la remarque précédente et la positivité de c_1 donnent $c_1 = \frac{1}{2}(-T_2 + \sqrt{T_2^2 + 4K_1})$ et $c_2 = T_2$, à condition de vérifier les inégalités :

$$\begin{cases} c_1 & < T_1(1) \\ T_2(c_1+T_2) & < K_2(2) \end{cases}$$

De $c_1(c_1+T_2)=K_1$, on trouve $T_2=\frac{K_1-c_1^2}{c_1}$, et en remplaçant dans (2), on obtient $P(c_1)<0$, où $P(c_1)=-c_1^2(K_1+K_2)+K_1^2$ est un polynôme de degré 2 en c_1, les racines de $P(c_1)$ sont T_1^p et $-T_1^p$. De plus comme le terme de second degré de P est négatif, on a : $T_1>T_1^p$.

Enfin en injectant la solution positive $c_1=\frac{1}{2}(-T_2+\sqrt{T_2^2+4K_1})$ dans (1), on trouve : $\frac{1}{2}(-T_2+\sqrt{T_2^2+4K_1})<T_1$, ce qui est équivalent à dire que : $T_1^2+T_1T_2-K_1>0$.

Pour résumer, il y a la présence d'un sel p1, pour :

$$\begin{cases} T_1 & > T_1^p \\ T_1^2+T_1T_2-K_1 & > 0 \end{cases}$$

Cas 4 :1 sel (p2) $=> p_2>0$ **et** $p_1=0$ Le système devient :

$$\begin{cases} K_1-(T_1-p_1)(T_1+T_2-p_1-p_2) & > 0 \\ K_2-(T_2-p_2)(T_1+T_2-p_1-p_2) & = 0 \\ p_1 & = 0 \\ p_2 & > 0 \end{cases}$$

En utilisant les expressions en fonction de c_1 et c_2, on obtient :

$$\begin{cases} c_1(c_1+c_2) & < K_1 \\ c_2(c_1+c_2) & = K_2 \\ c_1 & = T_1 \\ c_2 & < T_2 \end{cases}$$

On a un polynôme de degré 2 en c_2: $c_2(c_2+T_1)-K_2=0$ qui d'après la remarque précédente donne c_2. Donc, les solutions sont $c_2=\frac{1}{2}(-T_1+\sqrt{T_1^2+4K_2})$ et $c_1=T_1$, é condition de vérifier les inégalités :

$$\begin{cases} c_2 & < T_2(1) \\ T_1(c_2+T_1) & < K_1(2) \end{cases}$$

De $c_2(c_2+T_1)=K_2$, on trouve $T_1=\frac{K_2-c_2^2}{c_2}$, et en remplaçant dans (2), on obtient $P(c_2)<0$, où $P(c_2)=-c_2^2(K_1+K_2)+K_2^2$, les racines de $P(c_2)$ sont T_2^p et $-T_2^p$. De plus comme le terme de second degré de P est négatif , on a : $T_2>T_2^p$.

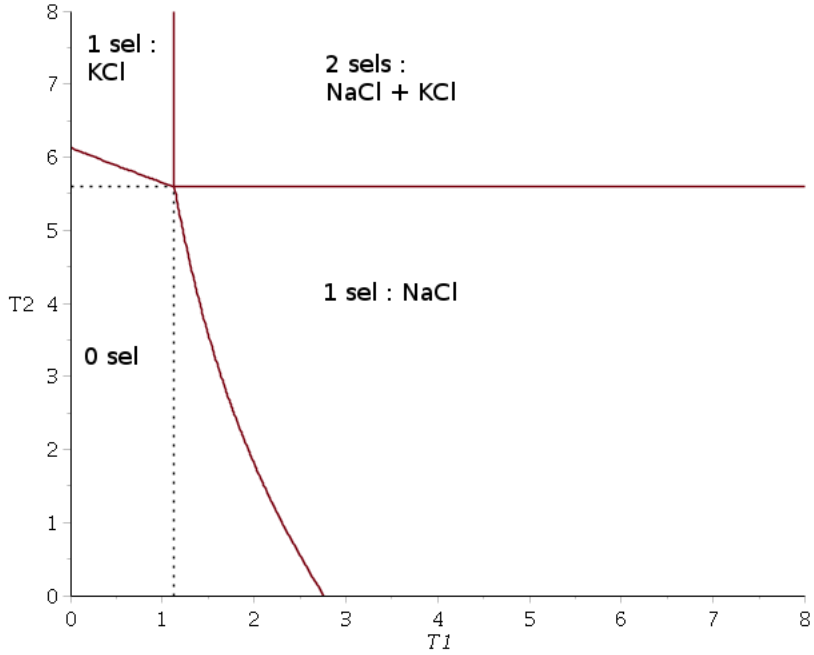
Enfin en injectant la solution positive $c_2=\frac{1}{2}(-T_1+\sqrt{T_1^2+4K_2})$ dans (1), on trouve : $\frac{1}{2}(-T_1+\sqrt{T_1^2+4K_2})<T_2$, ce qui est équivalent à dire que : $T_2^2+T_1T_2-K_2>0$.

Pour résumer, il y a la présence d'un sel p2, pour :

$$\begin{cases} T_2 & > T_2^p \\ T_2^2 + T_1 T_2 - K_1 & > 0 \end{cases}$$

7.5 Graphique

Dans le cas du $NaCl$ et du KCl, on utilise les constantes suivantes, K_1= 37.5837; K_2=7.6208. Grâce à l'étude faite précédemment on peut tracer le graphe de localisation du précipité en fonction des quantités initiales T_1 et T_2.



Présence de précipités. 2 sels : NaCl et KCl

8 Chimie avec 3 sels : Chlorure de Sodium, Chlorure de Potassium et Carnallite

On s'intéresse à la réaction de 3 précipités dans une solution, le Chlorure de Sodium ($NaCl$), le Chlorure de Potassium (KCl) et le Carnallite ($KMgCl3, 6H2o$)

$$\begin{aligned} Na^+ + Cl^- &\rightleftharpoons Nacl \\ K^+ + Cl^- &\rightleftharpoons Kcl \\ Mg^{++} + K^+ + 3Cl^- &\rightleftharpoons KMgCl3, 6H2o \end{aligned}$$

8.1 Variables

Soient c_1, c_2, c_3 et c_4 les variables de concentrations de Na^+, Cl^-, K^+ et Mg^{++} et p_1, p_2, p_3 les variables de concentrations des précipités $Nacl$, Kcl et $KMgCl3, 6H2o$ de plus K_1, K_2 et K_3 sont les produits de solubilités spécifiques aux réactions (constantes). Par interprétation physique des variables du problème les variables c et p sont positives.

8.2 Mise en équation

Lois d'action de masse minérale

$$\begin{aligned} min(K_1 - c_1c_4, p_1) &= 0 \\ min(K_2 - c_2c_4, p_2) &= 0 \\ min(K_3 - c_2c_3c_4^3, p_3) &= 0 \end{aligned}$$

Lois de conservation de masse

$$\begin{aligned} T_1 &= c_1 + p_1 \\ T_2 &= c_2 + p_2 + p_3 \\ T_3 &= c_3 + p_3 \\ T_4 &= c_4 + p_1 + p_2 + 3p_3 \end{aligned}$$

Neutralité électrique

$$T_4 = T_1 + T_2 + 2T_3$$

On élimine les variables $c3$, $p1$ et $p2$ pour obtenir :

$$\begin{aligned} min(K_1 - c_1c_4, T_1 - c_1) &= 0 \\ min(K_2 - c_2c_4, T_2 - c_2 - T_3 + c_3) &= 0 \\ min(K_3 - c_2c_3c_4^3, T_3 - c_3) &= 0 \end{aligned}$$

On peut aussi l'écrire comme un système avec 6 inégalités, où il y a toujours 3 égalités et 3 inégalités:

$$\begin{cases} c_1 & \leq T_1 \\ c_2 & \leq T_2 - T_3 + c_3 \\ c_3 & \leq T_3 \\ c_1c_4 & \leq K_1 \\ c_2c_4 & \leq K_2 \\ c_2c_3c_4^3 & \leq K_3 \end{cases} \tag{2}$$

avec $c_4 = c_1 + c_2 + 2c_3$.

8.3 Cas limite : toutes les inégalités sont actives

On cherche les éventuelles solutions au problème où toutes les inégalités du système sont actives, le système devient :

$$\begin{cases} c_1 & = T_1 \\ c_2 & = T_2 - T_3 + c_3 \\ c_3 & = T_3 \\ c_1 c_4 & = K_1 \\ c_2 c_4 & = K_2 \\ c_2 c_3 c_4^3 & = K_3 \end{cases}$$

On peut le simplifier :

$$\begin{cases} T_1 T_4 & = K_1 \\ T_2 T_4 & = K_2 \\ T_2 T_3 {T_4}^3 & = K_3 \end{cases}$$

D'après la neutralité électrique, on a :
$T_4 - T_1 - T_2 - 2T_3 = 0$
$\iff T_4 - \frac{K_1}{T_4} - \frac{K_2}{T_4} - 2\frac{K_3}{T_2 T_4^3} = 0$
$\iff T_4^3 - T_4(K_1 + K_2) - 2\frac{K_3}{K_2} = 0$

On obtient un polynôme de degré 3. On calcule un discriminant $D = \frac{K_3^2}{K_2^2} + \frac{1}{27}(-K_1 - K_2)^3$, on a deux cas :

- $D > 0$, alors il y a une racine réelle et 2 racines complexes et la racine réelle est : $T_4^p = (\frac{K_3}{K_2} + \sqrt{D})^{1/3} + (\frac{K_3}{K_2} - \sqrt{D})^{1/3}$.
- $D < 0$, alors il y a trois racines réelles distinctes :
  $T_4^p = \frac{2}{3}\sqrt{3K_1 + 3K_2} cos(\frac{1}{3} arccos(\frac{3K_3\sqrt{\frac{-3}{(-K_1-K_2)^3}}}{K_2}) + \frac{k\pi}{3})$, avec $k = 0$ ou 2 ou 4.
  On sait que $\frac{2}{3}\sqrt{3K_1 + 3K_2} > 0$, de plus $D < 0$, d'où
  $0 < 3\frac{K_3}{K_2}\sqrt{\frac{-3}{(-K_1-K_2)^3}} < 1$ et donc $1 < arccos(3\frac{K_3}{K_2}\sqrt{\frac{-3}{(-K_1-K_2)^3}}) < 1.5$, en utilisant cette connaissance dans l'expression $cos(\frac{1}{3} arccos(3\frac{K_3}{K_2}\sqrt{\frac{-3}{(-K_1-K_2)^3}}) + \frac{k\pi}{3})$, on voit que le seul k qui rend T_4^p positif est $k = 0$. D'où, $T_4^p = \frac{2}{3}\sqrt{3K_1 + 3K_2} cos(\frac{1}{3} arccos(3\frac{K_3}{K_2}\sqrt{\frac{-3}{(-K_1-K_2)^3}}))$.

on a ainsi :

$$T_1^p = \frac{K_1}{T_4^p},\ T_2^p = \frac{K_2}{T_4^p} \text{ et } T_3^p = \frac{K_3}{T_2^p (T_4^p)^3} = \frac{K_3}{K_2 (T_4^p)^2}$$

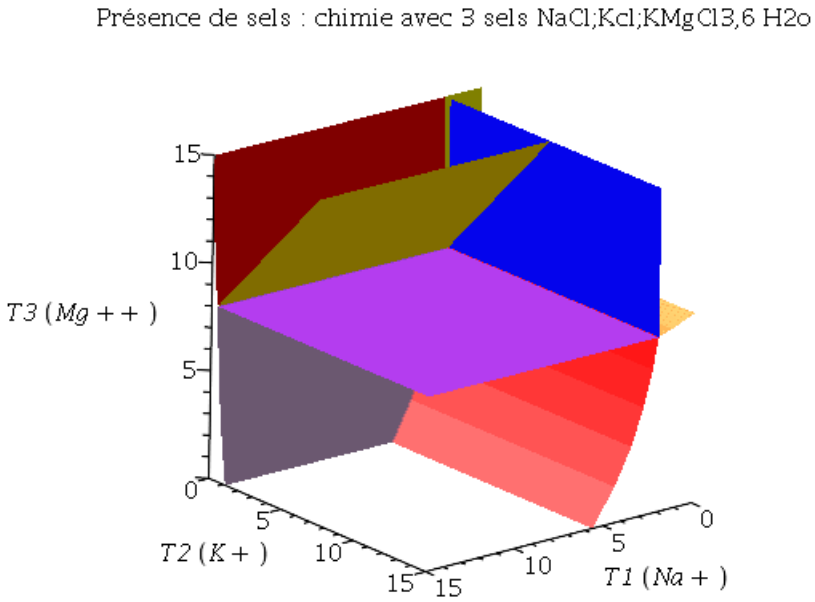
Dans ce cas le problème a une seule solution qui répond aux contraintes de positivité, $c_1 = T_1^p$, $c_2 = T_2^p$, $c_3 = T_3^p$ et $c_4 = T_4^p$.

8.4 Solutions

Il y a 8 façons de choisir une combinaison de 3 égalités et 3 inégalités dans le système (2), on procède de façon analogue à l'exemple précédent. Le détail du calcul est présenté en annexe.

8.5 Graphique

On trace les frontières entre les différentes phases en 3 dimensions avec comme variable T_1, T_2 et T_3, on utilise les constantes suivantes, K_1= 37.5837, K_2=7.6208 et K_3=2.1380e+04.



Sur ce graphique on distingue que chaque cas dans lequel se trouve le système est délimité par 3 hyperplans et d'autre part que tous les cas se rejoignent en un point c'est la solution du cas limite.

9 Généralisation de l'approche analytique

L'étude des 2 exemples précédents nous permet de mieux comprendre la généralisation d'une approche analytique. Cela reste une approche combinatoire donc limitée à des cas avec peu de minéraux, ce qui reste fréquent dans la pratique.

Comme dans les exemples précédents on utilise les lois de conservations et d'action de masse minérale pour obtenir un système d'équations/inéquations.

On choisit d'éliminer les variables c_i et donc d'avoir un système dépendant uniquement des variables p_i. Pour ce faire on utilise les lois de conservations :

$$c = T - E^t p$$

On obtient ainsi un système de $2N_p$ expressions avec N_p égalités et N_p inégalités

$$\begin{cases} K_i - \prod_{j=1}^{N_c} c_j(p,T)^{E_{ij}} & \geq 0 \\ p & \geq 0 \end{cases}, \ i = 1, \dots N_p \tag{3}$$

où le système (3) possède N_p inconnues p et les $N_c - 1$ paramètres T.

9.1 Point critique

On peut étudier le cas particulier où toutes les égalités sont actives

$$\begin{cases} K_i - \prod_{j=1}^{N_c} c_j(p,T)^{E_{ij}} & = 0 \\ p & = 0 \end{cases}, \ i = 1, \dots N_p \tag{4}$$

La solution de ce système est un point commun à tous les cas du système de chimie, c'est le **point critique**. On a N_p égalités qui nous donnent des valeurs pour p, en fait $p = 0$. Ensuite en remplaçant dans le le système on obtient N_p égalités qui dépendent de T et K ce qui va nous permettre de calculer N_p composantes de T. On pourra également utiliser la neutralité électrique pour éliminer un T_i. Le point critique du système dépend ainsi des constantes et de $N_c - N_p - 1$ paramètres T.

Si ce système n'admet pas de solution alors c'est que tous les minéraux ne peuvent être présent en même temps. On peut dans ce cas là chercher des sous-points critiques, c'est à dire les N_p point où $N_p - 1$ précipités réagissent simultanément.

9.2 Courbes critiques

Pour chaque réaction il y a deux cas possibles : soit le minéral est présent ou non. On considère N_p réactions, il y a donc 2^{N_p} cas possibles.

On va étudier tous les cas et dans chacun trouver les solutions, i.e des valeurs pour p et c en fonction de T et K, et trouver des inégalités qui ne dépendent que de T et K. Ces inégalités représentent les surfaces qui délimitent le cas dans lequel on se trouve, ce sont les **courbes critiques**. Une fois ces courbes critiques trouvées, lors de l'application numérique elles vont nous permettre de trouver dans quelle situation se trouve le système de chimie, i.e quels minéraux sont présents et en quelle quantité.

Pour chaque cas, on peut définir deux ensembles I et K tels que $I = \{i \in \{1, \dots, N_p\} |\ p_i = 0\}$ et $K = \{1, \dots, N_p\} \backslash I$. Le système devient

$$(Seq) \ : \ \begin{cases} K_i - \prod_{j=1}^{N_c} c_j(p,T)^{E_{ij}} & = 0 \ pour \ i \in K \\ p_i & = 0 \ pour \ i \in I \end{cases}$$

$$(Sineq) \ : \ \begin{cases} K_i - \prod_{j=1}^{N_c} c_j(p,T)^{E_{ij}} & \geq 0 \ pour \ i \in I \\ p_i & \geq 0 \ pour \ i \in K \end{cases}$$

Ensuite on utilise la démarche décrite dans l'algorithme suivant

Algorithm 1 Etude d'un cas

Soient (Seq) et (Sineq)

1. résoudre système polynomial (Seq) => $\forall i \in \{1..N_p\}$, $p_i = g(T)$
2. évaluer (Sineq) en $p_i = g(T)$
3. (Sineq) a N_p inégalités dépendantes uniquement de T=> courbes critiques qui délimite le cas.

9.3 Résolution du système polynomial

La difficulté de cette méthode est dans la résolution du système polynomial à N_p équations, c'est la partie qui demande le plus de temps de calcul et de mémoire. On remarque que dans la plupart des cas ce système est fortement simplifié par le fait qu'une partie des équations est triviale. On ne résout donc qu'une seule fois un système de N_p équations.

Dans le contexte d'étude d'un système polynomial, "résoudre" signifie dans un premier temps savoir le nombre de solutions du système. En effet si le système est dit "zéro-dimensionnel" alors il possède un nombre fini de solutions, si le système est de dimension positive alors il admet une infinité de solution et enfin il peut être inconsistant c'est à dire qu'il n'admet pas de solution. Afin de répondre à ces questions un outil très utilisé est les bases de Gröbner.

Algorithm 2 Caractéristiques d'un système polynomial

(S.0) Calculer la base de Gröbner du système polynomial pour un ordre lexicographique (en général grevlex)
(S.1) Est-ce que la base est réduite à 1 ?

- Oui, le système n'a pas de solution dans $\mathbb{C}$ (système inconsistant)
- Non, on continue

(S.2) Est-ce que pour chaque variable il y a un mônome de plus haut degré dans un élément de la base tel qu'il soit indépendant des autres variables ?

- Oui, le système est zéro-dimensionnel et possède donc un nombre fini de solutions
- Non, le système possède un nombre infini de solution dans $\mathbb{C}$

Dans le cas d'un système zéro-dimensionnel on peut calculer toutes les solutions. Bien souvent le calcul des ces solutions passe par une "triangularisation" du système, i.e transformé le système en un système de la forme :

$$\begin{cases} P_1(x_1, \dots x_n) & = 0 \\ \vdots & \vdots \\ P_{n-1}(x_1, x_2) & \\ P_n(x_1) & = 0 \end{cases}$$

La décomposition triangulaire peut aussi être réalisée pour des systèmes à dimensions positives.

Avec ce type de formulation les solutions dépendent très souvent de nombres algébriques, i.e des nombres qui sont représentés comme solution d'une équation algébrique. Cette remarque a deux conséquences sur l'évaluation numérique des solutions calculées :

1. Cela augmente l'aspect combinatoire de cette méthode, puisque en plus de devoir étudier tous les cas possibles il faut également considérer toutes les valeurs numériques réelles possibles pour chaque nombre algébrique.

2. Dans ce type de méthode la propagation des erreurs d'arrondi a un impact non-négligeable sur le résultat. Il est donc nécessaire de calculer les valeurs numériques des nombres algébriques avec une meilleure précision que la simple ou double précision. Plusieurs logiciels de calcul formel proposent ce type de calcul à précision arbitraire, on peut également mentionner la bibliothèque GMP (GNU Multiple Precision Arithmetic Library).

Dans le cas d'une résolution numérique, d'autres méthodes sont accessibles. La méthode de Newton peut être utilisé et on trouve plus de détails sur cette méthode dans la partie numérique de ce rapport. Lorsque le système possède autant d'équations que d'inconnues on peut utiliser une méthode d'homotopie.

Algorithm 3 Méthode d'homotopie pour résoudre un système polynomial

Soit $f_1 = 0, \dots f_n = 0$ le système polynomial que l'on souhaite résoudre.
(S.1) On calcul une borne sup (BS) du nombre de solutions (par exemple : lemme de Bézout)
(S.2) Soit $g_1 = 0, \dots g_n = 0$ un système qui possède BS solutions.
(S.3) Homotopie entre les 2 systèmes, par exemple :
on fait varier t de 0 à 1 pour

$$(1-t)g_1 + tf_1 = 0, \dots, (1-t)g_n + tf_n = 0$$

9.4 Application numérique

On a obtenu dans les parties précédentes le point critique du système et les courbes critiques délimitant chaque cas. On s'intéresse maintenant à l'application

numérique, c'est-à-dire comment obtenir la solution numérique du problème lorsque l'on donne des valeurs pour T.

Algorithm 4 Application numérique

Pour chaque cas possible **faire** :

1. donner une valeur numérique pour chaque nombre algébrique
2. **Si** une des composants de p ou c est négative pour la valeur du nombre algébrique
   (a) **Alors Si** il y a d'autres valeurs possibles pour les nombres algébrique **Alors** on essaye une autre et retour en (2)
   (b) **Sinon** on passe au cas suivant
3. **Sinon** on évalue (Sineq)
4. **Si** les inégalités de (Sineq) sont toutes vérifiées
   (a) **Alors** (Seq) donne la solution et on a fini (FIN).
   (b) **Sinon** on passe au cas suivant

FinPour

Pour résumer cette section, lorsque l'on souhaite étudier un système de chimie ne contenant que des réactions de précipitation-dissolution on suit les étapes suivantes :

1. On écrit le système comme un système à N_p équations et N_p inéquations
2. On cherche pour tous les cas possibles les courbes critiques (pour ce faire on a du calculer des valeurs de c et p en fonction de T pour chaque cas)
3. Enfin lorsque l'on souhaite obtenir la solution pour une valeur numérique du vecteur T : on cherche parmi tous les cas dans lequel on se trouve en utilisant les courbes critiques.

On va donc utiliser cette méthode pour un cas comportant 6 minéraux.

10 Chimie avec 6 sels : évolution de DMA en présence de 6 minéraux

On donne un exemple utilisé dans le cadre du transport réactif dans la littérature dans le contexte du DMA [19]. Un drainage minier acide (DMA) est une solution minérale acide qui apparaît dans certains types de mines ou dans le contexte du stockage de déchets miniers. Les problèmes du DMA sont principalement l'augmentation de l'acidité de l'eau (quasiment aucune espèce vivante ne vit

sous un pH de 5) et la formation d'acide sulfurique (H_2SO_4). Dans cet exemple on observe l'infiltration du DMA dans un milieu poreux fracturé. La première figure (Table 1) donne les composants qui sont présents dans les sources d'eaux de DMA et dans les eaux du milieu poreux initial. En dehors de la source sont également considérées des eaux de pluies qui agissent comme recharge.

Table 1
Initial background, as well as AMD source and non-source (natural recharge water) concentrations at surface for the conceptual model based on the tailings source at the Nickel Rim mine, Ontario Canada (Walter et al., 1994b).

Component	Initial background conc.[a] (mol/L)	Source conc. (mol/L)	Non-source conc. (mol/L)
Ca^{2+}	5.2e-03	1.08e-02	6.9e-03
Mg^{2+}	1.5e-03	9.69e-04	1.9e-03
Na^{+}	1.3e-03	1.39e-03	1.3e-03
K^{+}	6.6e-05	8.14e-04	6.7e-05
Cl^{-}	1.0e-03	1.58e-02	1.0e-03
CO_3^{2-}	2.4e-06	4.92e-04	3.9e-03
SO_4^{2-}	5.3e-03	5.00e-02	7.5e-03
Mn^{2+}	2.9e-05	7.84e-03	4.7e-05
$H_4SiO_4^0$	1.9e-03	2.08e-03	1.9e-03
Fe^{2+}	3.6e-05	3.06e-02	5.4e-05
Fe^{3+}	8.7e-17	2.00e-07	2.3e-08
Al^{3+}	6.6e-13	4.30e-03	1.3e-07
pH	7.2	4	7.0

[a] Small differences with Walter et al. (1994b) due to equilibration with modified background mineralogy.

On s'intéresse aux réactions de précipitation-dissolution du tableau 2, les constantes d'équilibres viennent de la base de MINTEQ/A2.

Table 2
Stoichiometric reactions and equilibrium constants K_k of all minerals in the DFN and EPM simulations (after Mayer et al., 2002; Molson et al., 2008).

Mineral	Reaction	Log K_k
Calcite	$CaCO_3 \Leftrightarrow Ca^{2+} + CO_3^{2-}$	−8.5
Siderite	$FeCO_3 \Leftrightarrow Fe^{2+} + CO_3^{2-}$	−10.5
Gibbsite	$Al(OH)_3 + 3H^+ \Leftrightarrow Al^{3+} + 3H_2O$	8.1
Gypsum	$CaSO_4 \cdot 2H_2O \Leftrightarrow Ca^{2+} + SO_4^{2-} + 2H_2O$	−4.6
Ferrihydrite	$Fe(OH)_3 + 3H^+ \Leftrightarrow Fe^{3+} + 3H_2O$	4.9
Quartz	$SiO_2\,(am) + 2H_2O \Leftrightarrow 3H_4SiO_4$	−2.7

Si on ne s'intéresse qu'aux phénomènes de précipitation-dissolution, on obtient un problème de 6 minéraux avec 8 composants (l'eau étant supposée en suffisamment grande quantité) . On peut également remarquer que l'on pourrait séparer ce problème en 3 sous-problèmes car les composants intervenant dans les réactions 3 et 5 n'interviennent pas dans les autres. La solution pour lier ces 3 problèmes nécessiterait de considérer des réactions de sorption.

Dans les calculs réalisés cette modification n'a pas été prise en compte, ce qui a pour conséquence de rendre les systèmes polynomiaux plus simple à résoudre mais le problème garde son aspect combinatoire.

10.1 Matrice stœchiométrique

D'après les tables précédentes, on a la matrice des coefficients stœchiométriques suivante :

$$\begin{pmatrix} E & c_1(H_4SiO_4) & c_2(H^+) & c_3(Fe^{3+}) & c_4(Al^{3+}) & c_5(Ca^{2+}) & c_6(SO_4^{2-}) & c_7(Fe^{2+}) & c_8(CO_3^{2-}) \\ p_6(Quartz) & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_3(Gibbsite) & 0 & -3 & 0 & 1 & 0 & 0 & 0 & 0 \\ p_5(Ferrihydrite) & 0 & -3 & 1 & 0 & 0 & 0 & 0 & 0 \\ p_1(Calcite) & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ p_2(Siderite) & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ p_4(Gypsum) & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

10.2 Mise en équation

Lois d'action de masse minérale

$$\begin{aligned} min(K_6 - c_1^3, p_6) &= 0 \\ min(K_3 - \frac{c_6}{c_2^3}, p_3) &= 0 \\ min(K_5 - \frac{c_3}{c_2^3}, p_5) &= 0 \\ min(K_1 - c_5c_8, p_1) &= 0 \\ min(K_2 - c_7c_8, p_2) &= 0 \\ min(K_4 - c_4c_5, p_4) &= 0 \end{aligned}$$

Lois de conservation de masse

$$\begin{aligned} T_1 &= c_1 + 3p_6 \\ T_2 &= c_2 - 3p_3 - 3p_5 \\ T_3 &= c_3 + p_5 \\ T_4 &= c_4 + p_4 \\ T_5 &= c_5 + p_1 + p_4 \\ T_6 &= c_6 + p_3 \\ T_7 &= c_7 + p_2 \\ T_8 &= c_8 + p_1 + p_2 \end{aligned}$$

La solution étant nécessairement électriquement neutre on a aussi la neutralité électrique qui donne une relation entre les T_i. On utilise les lois de conservation de masse pour obtenir les lois d'action de masse uniquement en fonction des précipités. Et on reformule comme le système à 6 équations/6 inéquations.

10.3 Cas limite : toutes les inégalités sont actives

Soit le système :

$$\begin{cases} K_1 - (T_5 + p_1 + p_4)(T_8 + p_1 + p_2) & = 0 \\ K_2 - (T_7 + p_2)(T_8 + p_1 + p_2) & = 0 \\ K_3 - \frac{T_6 + p_3}{(T_2 - 3p_3 - 3p_5)^3} & = 0 \\ K_4 - (T_4 + p_4)(T_5 + p_1 + p_4) & = 0 \\ K_5 - \frac{T_3 + p_5}{(T_2 - 3p_3 - 3p_5)^3} & = 0 \\ K_6 - (T_1 + 3p_6)^3 & = 0 \\ p_1 & = 0 \\ p_2 & = 0 \\ p_3 & = 0 \\ p_4 & = 0 \\ p_5 & = 0 \\ p_6 & = 0 \end{cases}$$

La solution à ce problème est $p_i = 0 \ \forall i$, ensuite on remplace dans les 6 premières égalités :

$$\begin{cases} K_1 - T_5 T_8 & = 0 \\ K_2 - T_7 T_8 & = 0 \\ K_3 - \frac{T_6}{(T_2)^3} & = 0 \\ K_4 - T_4 T_5 & = 0 \\ K_5 - \frac{T_3}{(T_2)^3} & = 0 \\ K_6 - (T_1)^3 & = 0 \end{cases}$$

On trouve des valeurs pour T, que l'on exprime en fonction de T_2 et T_7

$$\begin{cases} T_1 & = K_6^{1/3} \\ T_3 & = T_2^3 \\ T_4 & = \frac{K_4 K_2}{T_7 K_1} \\ T_5 & = \frac{T_7 K_1}{K_2} \\ T_6 & = T_2^3 K_3 \\ T_8 & = \frac{K_2}{T_7} \end{cases}$$

Avec la neutralité électrique on pourrait éliminer T_2 ou T_7 et obtenir le point critique en fonction d'une seule variable.

10.4 Solutions

Partie formelle : On utilise la méthode présentée dans la partie précédente pour générer automatiquement toutes les cas possibles et obtenir pour

chacun : la solution, i.e les valeurs de (c, p) par rapport aux T et K et les contraintes sur T qui délimitent ce cas. Il y a 6 minéraux donc 2^6=64 cas différents.

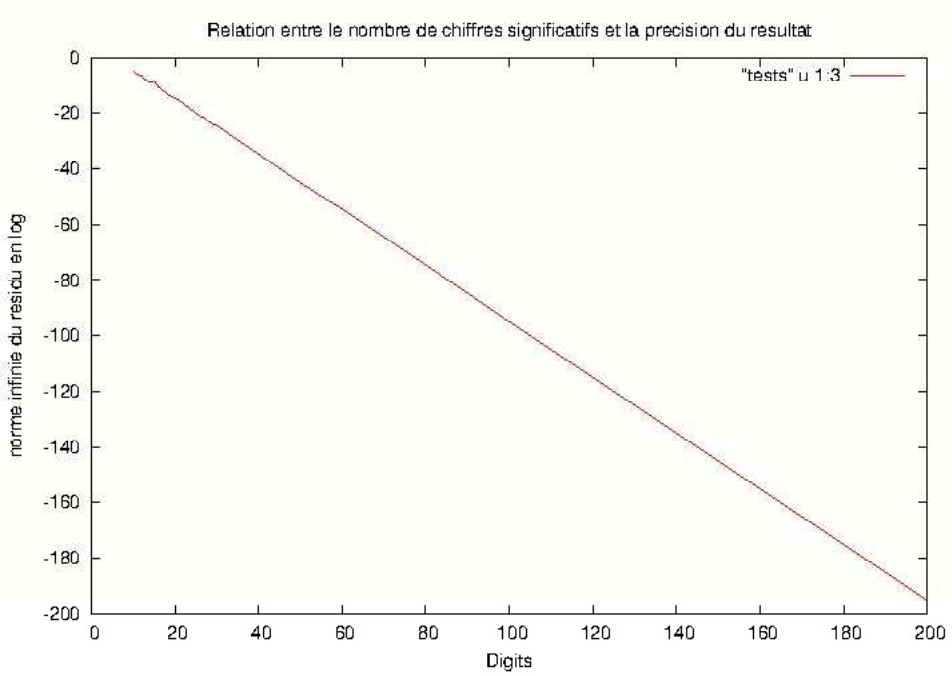
Pour cette exemple avec le logiciel de calcul formel MAPLE, on a :

étape	temps (en s)	mémoire (en MB)
Chargement bibliothèque de fonctions	0.96	33.93
Génération système de Chimie	<1	<1
Calcul des différentes phases	2.14	16.00
Total :	3.18	49.93

Les données de chargement de la bibliothèque de fonctions permettent d'obtenir un ordre de grandeur en temps et en mémoire. L'étape de génération du système de chimie est assez simple et relève essentiellement de la reformulation, elle est donc rapide. C'est la dernière étape qui supporte la partie combinatoire de l'étude il est donc logique que cela soit la plus coûteuse en temps et en mémoire. Ont ainsi été calculées les solutions pour toutes les phases du système en utilisant les nombres algébriques.

On remarque que pour ce cas à 6 sels, la difficulté combinatoire n'est pas si problématique.

Partie numérique : Dans un deuxième temps on souhaite évaluer ces solutions pour des valeurs numériques des constantes et des paramètres. La difficulté dans cette partie est l'évaluation des nombres algébriques. Pour ce faire on peut utiliser le logiciel de calcul formel MAPLE qui propose une extension de la double précision. Pour un T donné, on trace la courbe entre le nombre de digits utilisés sur MAPLE et la précision du résultat :



On voit que la relation est linéaire entre la précision du calcul et la précision du résultat.

11 Conclusion

Nous avons vu dans cette partie une méthode originale qui repose sur une approche combinatoire et utilise la structure polynomiale du système. L'intérêt principal de cette approche est de fournir des solutions du problème de façon exacte, en échange de quoi cette démarche nécessite de considérer tous les cas du systèmes et à chaque fois de résoudre un système polynomial.

L'application de cette méthode sur le cas 6 sels permet de tirer des premières conclusions sur ces difficultés. D'une part, on a vérifié que les résultats étaient précis dans la mesure où on peut utiliser une précision étendue. D'autre part, on a vu que la combinatoire n'est pas si problématique en temps de calcul et en mémoire, en effet le plus gros du travail qui consiste à trouver les courbes critiques étant réalisé une seule fois indépendamment du nombre d'application numérique. Enfin concernant la résolution des systèmes polynomiaux, l'exemple était suffisamment simple on a pu utiliser la fonction "solve" de MAPLE comme une boite noire, c'est un point qu'il faudra surveiller pour l'étude d'exemples plus complexes.

Maintenant on peut se demander les changements qui sont nécessaires pour l'intégration des réactions de sorption.

12 Ajout des réactions de sorption

On s'intéresse dans cette partie à l'intégration des réactions surfaciques au modèle où l'on a considéré que les réactions de précipitation-dissolution.

Dans l'étude analytique réalisée précédemment on avait :

$$\begin{cases} T - c - E^t p & = 0 \\ \varphi(\Pi(c), p) & = 0 \end{cases}$$

D'après le modèle défini en première partie on souhaite résoudre le système de la forme :

$$\begin{cases} T - T_c(c, s, p) & = 0 \\ W - T_s(c, s) & = 0 \\ \alpha_i = K_\alpha \prod_{l=1}^{N_c} c_l^{S_{i,l}} & i = 1, ..N_\alpha \\ \beta_j = K_\beta \prod_{l=1}^{N_c} c_l^{A_{j,l}} \prod_{m=1}^{N_s} s_m^{B_{j,m}} & j = 1, ..N_\beta \\ \varphi(\Pi(c), p) & = 0 \end{cases}$$

où

$$\begin{cases} T_c(c, s, p) & = c + S^t \alpha(c) + A^t \beta(c, s) + E^t p \\ T_s(c, s) & = s + B^t \beta(c, s) \end{cases}$$

Dans l'étude analytique que l'on a réalisé on exprime l'équation $\varphi(\Pi(c), p) = 0$ comme un système à N_p équations et N_p inéquations que l'on exprimait en fonction de N_p variables. On choisissait d'exprimer ces expressions en fonction des variables p car on avait une relation direct entre c et p.

$$c = T - E^t p$$

Or dans le cas avec les réactions de sorption cette relation entre les c et les p n'est plus aussi simple.

$$\begin{cases} T & = c + S^t \alpha(c) + A^t \beta(c, s) + E^t p \\ W & = s + B^t \beta(c, s) \end{cases}$$

On obtient un système polynomial de $N_c + N_s$ équations qu'il faut résoudre si on souhaite obtenir les valeurs de c en fonction de p.

En pratique il faut choisir les N_p variables (pas obligatoirement les p) de façon à ce que ce système polynomial soit le plus simple possible. On peut déjà obtenir des valeurs pour p en fonction de c et s en résolvant un système linéaire

$$E^t p = b(c, s)$$

Il faudra à l'avenir considérer des exemples pour voir si en pratique il est difficile de trouver N_p variables qui rendent le système polynomial simple à résoudre.

Part III
Étude numérique

Une des difficultés du système de chimie est la présence d'un problème de complémentarité. Dans le modèle de chimie on a parlé d'un (PC) sous la forme :

$$(PC)\ \varphi(a,b) = 0 \iff ab = 0\ et\ a \geq 0\ b \geq 0$$

Mathématiquement, le problème de complémentarité consiste à trouver un vecteur $p \geq 0$ tel que $F(p)\ \geq 0$ et $p^T F(p) = 0$, aussi écrit sous forme contracté :

$$0 \leq p \perp F(p) \geq 0$$

On l'appelle problème de complémentarité linéaire si on peut écrire $F(p)$ comme $Mp + q$, où M est une matrice réel et q un vecteur réel. Dans le cas contraire on a un problème de complémentarité non-linéaire, ce qui est le cas du système de chimie que l'on considère dans ce rapport.

Pour résoudre ce type de problème une méthode très utilisée est de reformuler ce problème grâce à une C-fonction. On a déjà vu dans la première partie, que l'on peut utiliser la C-fonction min (qui prend le minimum composante par composante) $\varphi_{min}(a,b) = min(a,b)$.

$$(PC) \iff min(a,b) = 0$$

Grâce à cette technique de reformulation, on a un système de la forme :

$$G(c,p) = 0$$

Il est clair que la fonction $min(a,b)$ n'est pas continûment différentiable, par contre elle est globalement continûment lipschitzienne et continûment différentiable pour $a \neq b$. On peut même montrer qu'elle est fortement semi-lisse, voir [4] les définitions et la démonstration. De plus dans [4], on a aussi les propriétés qui font que le système $G(c,p)$ est fortement semi-lisse, c'est donc une méthode de Newton semi-lisse qui va être utilisée.

Une autre difficulté est d'ajouter des contraintes de positivité aux variables c_i. En effet comme les variables c_i ne sont plus exprimées par des log elles peuvent devenir négatives, il faut donc forcer les variables c_i à rester positive. Nous verrons par la suite comment prendre en compte ces contraintes dans une méthode de Newton.

13 Étude du jacobien généralisé

Soit le système de la forme

$$G(c,p) = \begin{cases} T - c - E^t p = 0, \\ min(\, K_i - \prod_{j=1}^{N_c} c_j^{E_{ij}}, p_i \,) = 0, \quad i = 1, ..., N_p \end{cases}$$

Ce système n'est pas continûment différentiable par contre il est fortement semi-lisse et on peut calculer le jacobien généralisé. Afin de calculer la dérivée de la fonction min on utilise une stratégie d'ensemble actif/inactif. Soit $A = \{i \in \{1, ..., N_p\} \mid K_i - \prod_{j=1}^{N_c} c_j^{E_{ij}} \geq p_i\}$ et $I = \{1, \ldots, N_p\}/\mathrm{A}$, on a donc le système :

$$\begin{cases} T - c - E^t p = 0 \\ K_i - \prod_{j=1}^{N_c} c_j^{E_{ij}} = 0 & i \in A \\ p_i = 0 & i \in I \end{cases}$$

La matrice jacobienne par rapport aux variables (c, p) du système est ainsi de la forme :

$$J(c,p) = \begin{pmatrix} -I_{Nc} & -E_A^t & -E_I^t \\ X_1 & 0 & 0 \\ 0 & 0 & I_{N_I} \end{pmatrix}$$

où ont été permutées des lignes pour avoir les termes actifs et inactifs dans les mêmes blocs. A chaque itération de la méthode de Newton il faut résoudre un système linéaire. Pour être sûr que ce système a toujours une unique solution il faut montrer que le déterminant de la jacobienne est non nul quels que soient c et p, i.e $det(\, J(c,p)) \neq 0$.

On peut décomposer la matrice jacobienne comme le produit d'une matrice triangulaire inférieure à diagonale composée de matrice identité et d'une matrice triangulaire supérieure, on a :

$$J = LU$$

$$\begin{pmatrix} -I_{Nc} & -E_A^t & -E_I^t \\ X_1 & 0 & 0 \\ 0 & 0 & I_{N_I} \end{pmatrix} = \begin{pmatrix} I_{Nc} & 0 & 0 \\ -X_1 & I_{N_A} & 0 \\ 0 & 0 & I_{N_I} \end{pmatrix} \begin{pmatrix} -I_{Nc} & -E_A^t & -E_I^t \\ 0 & -X_1 E_A^t & -X_1 E_I^t \\ 0 & 0 & I_{N_I} \end{pmatrix}$$

En passant au déterminant :

$$det(J) = det(LU) = det(L)det(U) = det(X_1 E_A^t)$$

Considérons maintenant le bloc X_1, on a :

$$X_1 = \frac{\partial}{\partial c}(\, K_i - \prod_{j=1}^{N_c} c_j^{E_{ij}})_{i \in A}$$

Soit l'hypothèse $H_1 := \{c > 0\}$, sous cette hypothèse X_1 devient :

$$X_1 = (\frac{E_{A,ij}}{c_j}\prod c_l^{E_{A,il}})_{i\in A\ 1\leq j\leq Nc} = diag(\,(\prod c_l^{E_{A,il}})_{i\in A})\,E\,diag(\frac{1}{c})$$

où diag est la matrice diagonale, et $\frac{1}{c} = [\frac{1}{c_1}, ..., \frac{1}{c_{Nc}}]$. Le déterminant devient :

$$\begin{aligned} det(X_1E_A^t) &= det(\,diag(\,(\prod c_l^{E_{A,il}})_{i\in A})E_A\,diag(\frac{1}{c})\,E_A^t) \\ &= det(\,diag(\,(\prod c_l^{E_{A,il}})_{i\in A}))\,det(\,E_Adiag(\frac{1}{c})E_A^t) \\ &= \prod_i\prod_j c_l^{E_{A,il}}det(E_Adiag(\frac{1}{c})E_A^t) \end{aligned}$$

Or d'après H_1 on voit que $\prod_i\prod_j c_l^{E_{A,il}} \neq 0$.

$$det(E_Adiag((\frac{1}{c}))E_A^t) =?$$

Dans un premier temps on remarque que :

$$M = E_Adiag((\frac{1}{c}))E_A^t = (E_Adiag((\frac{1}{c}))E_A^t)^t$$

De plus on peut écrire M comme :

$$M = E_Adiag(\frac{1}{\sqrt{c}})diag(\frac{1}{\sqrt{c}})E_A^t = (E_Adiag(\frac{1}{\sqrt{c}}))(E_Adiag(\frac{1}{\sqrt{c}}))^t$$

Ensuite E_A est inversible par hypothèse sur l'indépendance des lignes dans le système chimique, et $diag(\frac{1}{c})$ est inversible aussi d'où $E_Adiag(\frac{1}{\sqrt{c}})$ est inversible.

On en conclut que M est symétrique définie positive, et ainsi :

$$det(E_Adiag((\frac{1}{c}))E_A^t) \neq 0$$

Enfin en remontant notre raisonnement sous l'hypothèse H_1, on a $det(J(c,p)) \neq 0$ et la matrice jacobienne est non-singulière.

Remarque : L'hypothèse H_1 n'est pas injustifiée, en effet les concentrations ne peuvent être négatives car le système chimique n'aurait pas de sens physique. Par contre il faut étudier le déterminant de $X_1E_A^t$ dans le cas où les concentrations sont nulles.

Remarque 2 : Est-il possible de généraliser cette démonstration à la présence des réactions de sorption ?

14 Méthode de Newton semi-lisse

Dans cette section nous allons rappeler les éléments essentiels à l'étude numérique du système de chimie concernant la méthode de Newton semi-lisse et les méthodes de globalisation. Dans le cas du système de chimie G que l'on étudie $n = N_c + N_p$ et le vecteur $x \in \mathbb{R}^{N_c+N_p} = \begin{pmatrix} c \\ p \end{pmatrix}$.

La méthode de Newton permet de résoudre le problème non-linéaire $G(x) = 0$, $G : \mathbb{R}^n \to \mathbb{R}^n$ de n équations à n inconnues. Soit $G'(x_k)$ le jacobien de G au point x_k. On a d'abord le résultat de convergence classique suivant :

Théorème (convergence locale) : S'il existe x^* une solution de $G(x) = 0$, $G' : \Omega \to \mathbb{R}^{n\times n}$ est continûment lipschitzien autour de x^* et $G'(x^*)$ est inversible alors si x_0 est suffisamment proche de x^* il existe une séquence de Newton (i.e $G'(x_n)$ non-singulière $\forall n > 0$) qui converge vers x^*, de plus il existe $K > 0$ tel que pour n suffisamment large : $||e_{n+1}|| \leq K||e_n||^2$ (convergence q-quadratique)

On rappelle l'algorithme utilisé pour la méthode de Newton :

Algorithm 5 Méthode Newton semi-lisse

(S.0) Initialisation
Soit x^0 l'itérée initiale et τ la précision
(S.1) Critère d'arrêt
$||G(x_k)|| < \tau$
(S.2) Direction de descente
Trouver d tel que $G'(x_k)d = -G(x_k)$
(S.3) Mise à jour
$-x_{k+1} = x_k + d$
retour en (S.1)

Comme on le voit la méthode de Newton est une méthode locale or avoir un bon point initial n'est en pratique pas toujours réalisable. Néanmoins, on peut utiliser des techniques de globalisation, c'est-à-dire qu'il n'y aura pas de convergence quadratique globale mais que quel que soit le point initial choisi on va converger vers un voisinage de la solution où il y aura une convergence quadratique. Il existe essentiellement 2 classes de méthodes de globalisation : les méthodes de recherche en ligne (line search) et les méthodes de région de confiance (trust-region). Dans chacun des cas on a besoin d'une fonction objectif que l'on va chercher à minimiser. En général c'est la fonction suivante :

$$\psi(x) = ||G(x)||_2^2$$

Line search (Recherche en ligne) Plutôt que de considérer la simple direction de Newton comme avant, on va nommer d la direction de Newton et λ le pas de la direction de Newton, ainsi soit $d = -G'(x_n)^{-1}G(x_n)$ et :

$$x_{k+1} = x_k + \lambda d$$

Cette méthode s'appelle recherche en ligne car on cherche une baisse de $\psi(x)$ le long du segment $[x_n, x_n + d]$.

Il existe plusieurs techniques pour déterminer le pas. Une première est de choisir $\lambda = 2^{-m}$, où m est le plus petit entier positif tel que :

$$||G(x_n + 2^{-m}s)|| < (1 - \alpha 2^{-m})||G(x_{n)}||$$

avec $\alpha \in [0, 1]$. C'est la règle d'Armijo.

Certains cas réagissent bien à la réduction de la règle d'Armijo, pour d'autres cas il faudrait beaucoup de réductions de ce type. Pour pallier à ce problème on utilise un modèle quadratique basé sur les trois plus récentes valeurs de λ qui intervient si après 2 réductions il n'y a pas eu de baisse suffisante de

$$\Phi(\lambda) = ||G(x_n + \lambda d)||^2$$

Le λ suivant est le minimum du modèle quadratique avec comme vérification que la réduction en λ est d'au moins un facteur 2 et au plus un facteur 10. Ainsi l'algorithme génère une séquence de candidats λ_m avec λ_0=1 et $\frac{1}{10} \leq \lambda_{m+1}/\lambda_m \leq \frac{1}{2}$. La recherche en ligne se termine avec le plus petit $m \geq 0$ tel que :

$$||G(x_n + \lambda_m d)|| < (1 - \alpha\lambda_m)||G(x_n)||$$

C'est cette méthode qui est implémentée dans les codes de [12, 11].

Trust-région La méthode de région de confiance est duale à la recherche en ligne. Dans le cas de la région de confiance, d'abord est choisie la taille de la région puis la direction tandis que dans la recherche en ligne d'abord on a la direction ensuite le pas. Pour en savoir plus, voir [31].

15 Prise en compte des contraintes

Dans cette partie plusieurs méthodes de prise en compte des contraintes de positivité des variables sur une méthode de Newton ont été essayées. Tout d'abord une méthode active/set proposée dans la littérature, puis une méthode de modification de pas de la méthode présente dans le solveur KINSOL et enfin une méthode de pénalisation de la fonction objectif.

15.1 Kanzow Active-Set

Pour respecter le sens physique du système de chimie à chaque itération de la méthode de Newton semi-lisse on utilise une méthode hybride qui génère uniquement des itérées réalisables. Dans notre cas les contraintes étant uniquement des contraintes de positivité, on est dans le cas particulier des "box constrained".

Un premier algorithme respectant ces contraintes est celui développé dans [22] et [23].

$$F(x) = 0,\ x \in [l, u]$$

On utilise comme fonction de descente :

$$min\ f(x) = \frac{1}{2}||F(x)||^2\ tel\ que\ x \in [l, u]$$

On présente ici la méthode de façon générale pour des contraintes du type $x \in [l, u]$, pour le cas du système de chimie on a un cas particulier de ce genre de problème, i.e :

$$G(c, p) = 0,\ (c, p) \in [0, +\infty[$$

et la fonction de descente :

$$min\ g(x) = \frac{1}{2}||G(x)||^2\ tel\ que\ (c, p) \in [0, +\infty[$$

Algorithm 6 Active Set type Newton Method

(S.0) (Initialisation)
Soit $x^0 \in [l, u]$, $\beta \in (0,1)$, $\gamma \in (0,1)$, s>0, $\delta > 0$, h>0, et k:=0
(S.1) (Critère d'arrêt)
Si x^k est un point stationnaire : STOP
(S.2) (Active Set Strategy)
Soit $\delta_k := min(\delta, h\sqrt{||F(x^k)||})$
et
$\mathcal{A}_k := \{i | x_i^k - l_i \leq \delta_k \ ou \ u_i - x_i^k \leq \delta_k\}$,
$\mathcal{I}_k := \{1, ..., n\} \backslash \mathcal{A}_k$
(S.3) (Direction de Newton)
Soit F'(x) la jacobienne de F(x)
Pour $i \in \mathcal{A}_k$, alors
$d_i^k := \begin{cases} l_i - x_i^k & si \ x_i^k - l_i \leq \delta_k \\ u_i - x_i^k & si \ u_i - x_i^k \leq \delta_k \end{cases}$
puis on résout le système linéaire, pour obtenir les composantes de d_k pour $i \in \mathcal{I}_k$
$F'_{\mathcal{I}_k \mathcal{I}_k}(x^k) d_{\mathcal{I}_k} = -F(x^k) - F'_{\mathcal{I}_k \mathcal{A}_k}(x^k) d^k_{\mathcal{A}_k}$
(S.4) (Feasibility Preserving Safeguard)
$\tau_k := sup\{\tau > 0 \ x_i^k + \tau d_i^k \in [l_i, u_i] \ \forall i \in \mathcal{I}_k\}$
$\tau_k := min\{1, \tau_k\}$
(S.5) (Computation of New Iterate)
Si d^k a pu être calculé et satisfait la condition $f(x^k + \tau_k d^k) \leq \gamma f(x^k)$
alors $x^{k+1} := x^k + \tau_k d^k$
Sinon on utilise une étape de gradient projeté, soit :
$x^{k+1} := x^k(t_k)$oé $x^k(t) := [x^k - t\nabla f(x^k)]_+$ et $t_k = max\{s\beta^l, l = 0, 1, 2...\}$ tel que
$f(x^k(t_k)) \leq f(x^k) - \sigma \nabla f(x^k)^T (x^k - x^k(t_k))$
(S.6) (Update)
k:=k+1

Algorithme Maintenant quelques remarques sur l'algorithme et son implémentation.

Comme critère d'arrêt sont utilisés trois critères : en premier lieu si le résidu est inférieur à la précision demandée on arrête et l'algorithme a réussi, ensuite si la différence relative du résidu entre les 2 dernières itérations on arrête et l'algorithme stagne enfin on place un critère d'itération maximum.

Le δ doit être suffisamment petit pour ne pas être vérifié par les deux "bords" simultanément, donc :

$$\delta < \frac{1}{2} min_{i=1,...n} |u_i - l_i|$$

Dans (S.4), on choisit le pas maximum qui ne sort pas des contraintes, une

version simplifiée de l'expression utilisée est :

$$\tau_k := min\left\{min_{\{i\in\mathcal{I}_k|d_i^k<0\}}\frac{l_i-x_i^k}{d_i^k}, min_{\{i\in\mathcal{I}_k|d_i^k>0\}}\frac{u_i-x_i^k}{d_i^k}\right\} \ et \ \tau_k = min\{1,\tau_k]$$

Dans (S.5), on cherche à accepter le nouvel itéré calculé par la direction de Newton. En pratique, on utilise une méthode Armijo de recherche en ligne pour avoir plus de chance d'accepter cet itéré.

Si l'itéré n'est pas accepté, on utilise une méthode de gradient projeté. On peut noter que la direction du gradient a une expression simple ici, en effet :

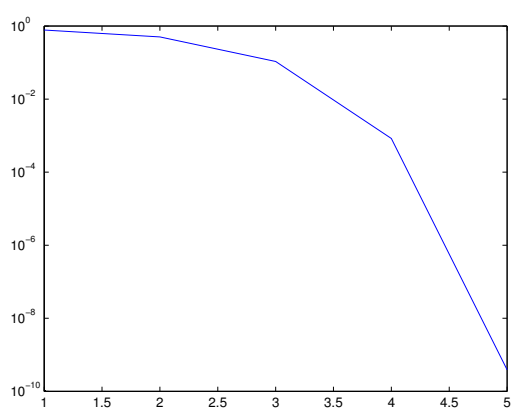
$$\nabla f(x^k) := (F'(x^k))^T f(x^k)$$

D'autre part, la projection sur une box [l,u] est aussi très simple, en effet $\forall z$ on prend :

$$z_+ := max(l, min(x, u))$$

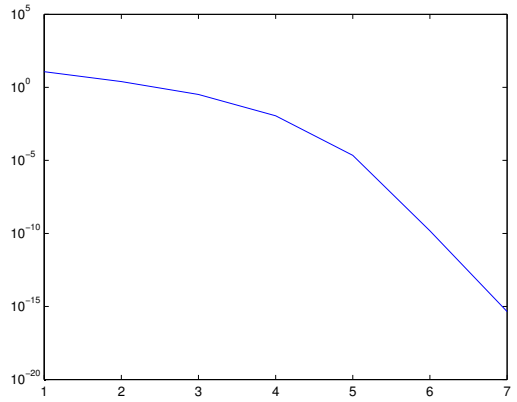
Résultats numériques : On présente dans un premier temps 2 exemples simples pour montrer que la méthode conserve les qualités de convergence d'une méthode de Newton. Puis sont montrés deux exemples dans notre contexte le cas test "Easy Test Case" du benchmark MoMas.

Exemple 1 : $F(x) = arctan(x)$



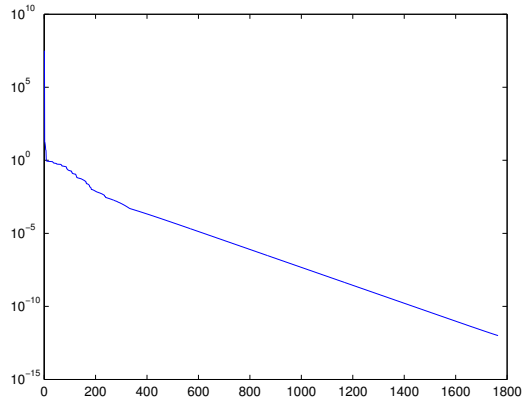
La méthode a une convergence d'ordre 3 qui correspond à l'ordre obtenu si on utilise uniquement une méthode de Newton avec line search.

Exemple 2 : $F(x) = \begin{cases} x^2+y^2-2 \\ x^2-y^2-1 \end{cases}$



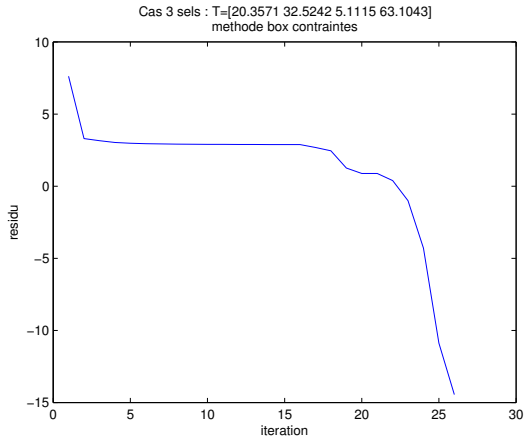
La méthode a une convergence d'ordre 2 qui correspond à l'ordre obtenu si on utilise uniquement une méthode de Newton avec line search.

Exemple 3 : MoMas Benchmark "EasyTestCase" (voir Annexe 2), où on a réduit l'écart entre les constantes pour rendre le problème légèrement moins raide (K5=10 au lieu de K5=1e35)



On voit que la méthode fait des progrès rapide mais ensuite devient "lente". Dans les faits l'utilisation très fréquente du gradient comme direction de descente ralenti l'ordre de convergence.

Exemple 4 : cas 3 sels



On observe une descente rapide sur les 10 premières itérations, puis une convergence linéaire, pour atteindre la précision demandée.

Conclusion : Cette méthode ne semble pas être réellement adaptée au problème avec un min, en effet dans les cas pratique de l'étude chimique il faut beaucoup d'itérations, essentiellement causées par l'utilisation trop fréquente de la direction de gradient comme direction de descente.

15.2 Modification du pas

Le solveur KINSOL de la suite SUNDIALS [2] propose une méthode de Newton avec technique de globalisation line search pour résoudre les problèmes non-linéaires. D'autre part il proposer la prise en compte de contraintes supplémentaires de la forme $u_i \geq 0$ ou $u_i \leq 0$ où u est le vecteur solution du problème. KINSOL va réduire le pas de la méthode pour être certain qu'aucune contrainte ne soit violée. Si un itérée de Newton ne respecte pas une contrainte, le pas maximum le long de la direction de Newton qui satisfait toutes les contraintes est trouvé. Et dans la méthode line search on ne cherche plus un pas entre $[0, 1]$ mais entre $[0, \lambda_{max}]$. On va utiliser cette technique pour les contraintes de positivité sur c.

Autrement dit, si d est la direction de Newton et λ le pas de la direction de Newton, ainsi soit $d = -G'(x_n)^{-1}G(x_n)$ et :

$$x_{k+1} = x_k + \lambda d$$

On cherche une baisse de $\psi(x) = ||G(x)||_2^2$ le long du segment $[x_n, x_n + \lambda_{max} d]$.

15.3 Newton avec pénalisation de la fonction objectif

Dans cette méthode on utilise une méthode de Newton avec line search où c'est la fonction objectif qui va être pénalisée pour respecter les contraintes. La fonction objectif que l'on souhaite minimiser devient

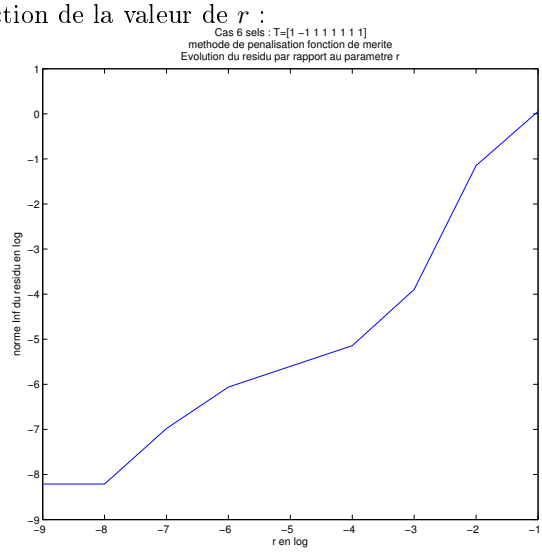
$$\psi(x) = ||G(x)||_2^2 + \frac{1}{N_c^2}(e\ exp(-c/r))$$

où e est le vecteur de $(\mathbb{R}^{N_c})^T$ qui ne contient que des 1 et r est un paramètre qui tend vers 0. Observons le comportement du terme $exp(\frac{-c_i}{r})$ en fonction de c_i :

- $c_i > 0$; alors comme r$\rightarrow 0$, $\frac{-c_i}{r} \rightarrow -\infty$ et donc $exp(\frac{-c_i}{r}) \rightarrow 0$
- $c_i = 0$; $exp(\frac{-c_i}{r}) = 1$
- $c_i < 0$; alors comme r$\rightarrow 0$, $\frac{-c_i}{r} \rightarrow +\infty$ et donc $exp(\frac{-c_i}{r}) \rightarrow +\infty$

Ensuite on utilise l'algorithme de Newton semi-lisse et dans la procédure de globalisation line search on utilise la fonction objectif modifiée. Le choix de r dépend de la difficulté du problème, il faut faire plusieurs essais pour trouver la bonne valeur et choisir un r suffisamment petit tel qu'il n'y ait pas d'itérée négatif.

Résultats numériques évolution de la valeur du résidu en norme infinie en fonction de la valeur de r :

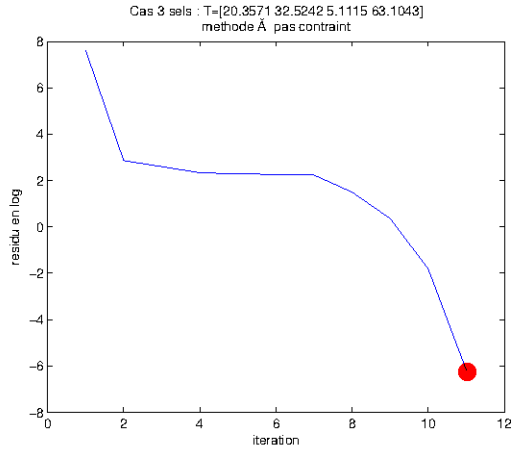


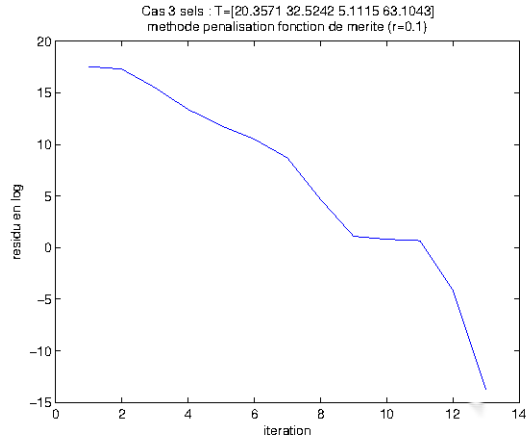
La baisse du paramètre r fait de progrès jusqu'à un plateau. Pour chaque type de problème il faut faire une procédure d'essai pour trouver la valeur de r adéquate, dans ce cas 6 sels c'est $r = 1e - 8$.

15.4 Comparaison des méthodes

On reprend les exemples 3 sels et 6 sels précédents pour comparer les 2 méthodes de prise en compte des contraintes que l'on a sélectionné. On choisit comme point initial le vecteur des conditions initiales, i.e $c_0 = T$ et $p_0 = 0$

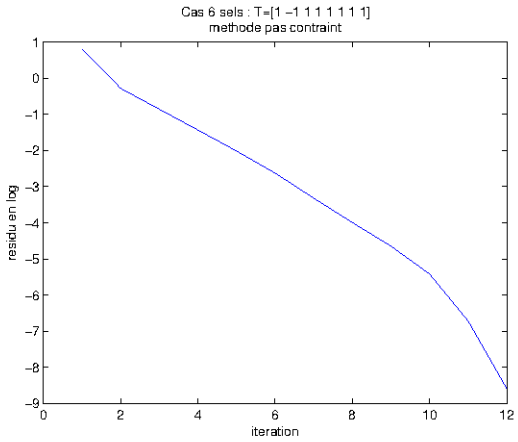
Comparaison rapidité et précision : Cas 3 sels T=[20.3571 32.5242 5.1115 63.1043]

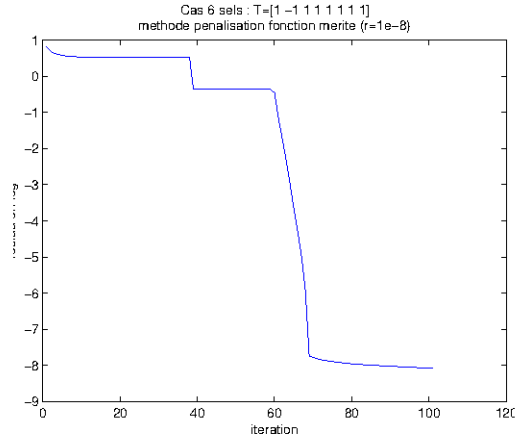




Dans les deux cas on observe le même phénomène en 3 phases. En premier le résidu diminue rapidement, puis il y a un phénomène de palier et enfin une baisse rapide. On voit que dans le cas de la méthode à pas contraint les itérations s'arrêtent, en effet avec la méthode line search on s'arrête si le long de la direction de descente calculée on ne peut plus améliorer la fonction objectif.

Cas 6 sels T=[1 -1 1 1 1 1 1 1]





A nouveau on observe le même phénomène de palier que dans le cas 3 sels, mais cette fois si la méthode à pas contraint à été beaucoup plus rapide.

Au vu de ces résultats la méthode à pas contraint semble plus rapide, il reste à voir si elle termine souvent sur un échec ce qui a pour conséquence de perdre en robustesse et en précision.

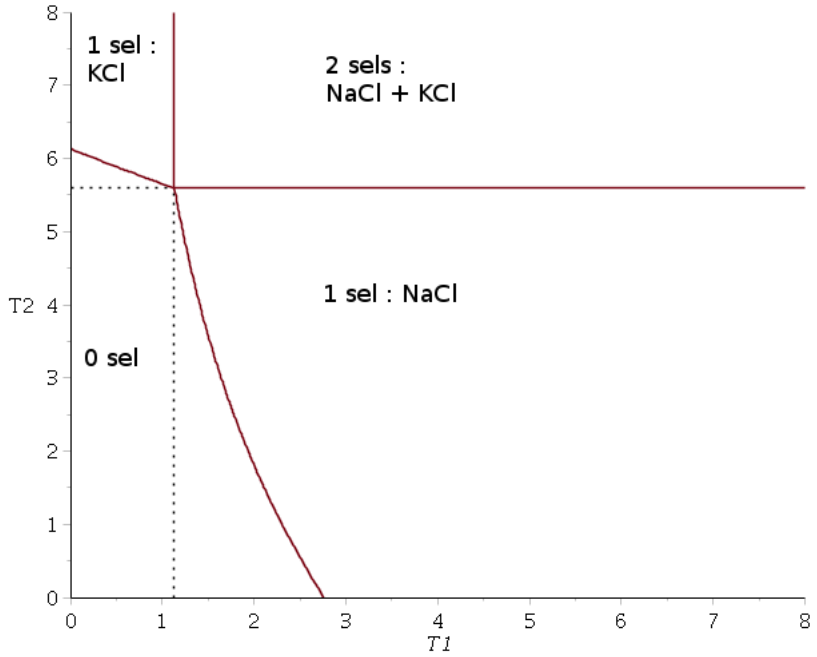
Pourcentage d'échec des méthodes : On calcul un pourcentage où la méthode échoue, i.e pour une méthode line search on a une direction de descente mais on n'arrive pas à baisser le résidu le long de cette direction. On réalise 10 000 tirages aléatoires de valeur de T pour le cas 3 sels. Pour la méthode pénalisation on a 2.5% d'échec et pour la méthode à pas contraint 40%.

En conclusion il semble que la méthode avec pénalisation de la fonction objectif soit préférable car même si elle est plus lente elle échoue moins souvent et ainsi en moyenne offre une meilleure précision.

Enfin concernant le comportement en 3 phases de la méthode, une interprétation possible est de dire que l'on part d'un point initial que l'on peut rapidement améliorer jusqu'à un certain palier ensuite la méthode va de cas en cas pour trouver la bonne zone et ainsi le voisinage de la solution qui assure une convergence rapide pour terminer. On peut supposer qu'en améliorant le point initial, on pourrait éviter ce phénomène de palier et gagner du temps de calcul.

16 Choix du point initial

On essaye de trouver un "bon" point initial pour éviter un phénomène de palier. On peut utiliser notre connaissance du système chimique acquise dans la partie analytique pour nous aider. On a vu que chaque cas est délimité par des surfaces et ont toutes en commun le point critique. Une approximation de cette observation serait de dire que chaque cas est délimité par des plans, on a ainsi les courbes critiques et en fonction des valeurs de T on peut dire si un minéral est présent ou non. Ainsi on peut essayer de trouver un point initial qui est dans le bon cas.

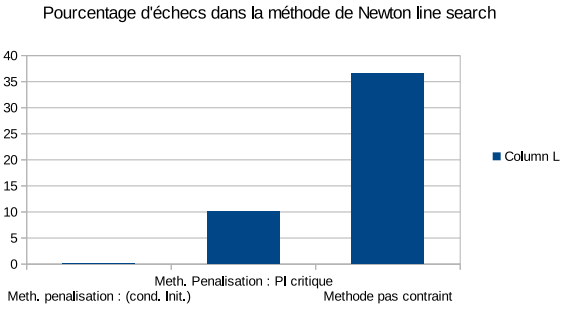


Présence de précipités. 2 sels : NaCl et KCl

Par exemple pour le cas 2 sels, on calcul le point critique. Ensuite on réalise une approximation des courbes critiques par des droites. Les 4 courbes critiques obtenues ainsi sont les 2 en pointillés et les 2 qui délimitent le cas où les 2 sels sont présents. Une fois que l'on a identifié le cas dans lequel on se trouve, on peut choisir un point initial dans ce cas.

Résultats numériques avec ce point initial :

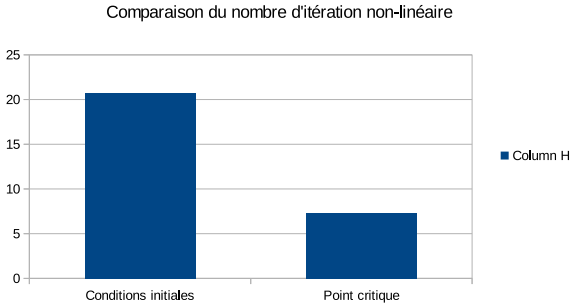
- On calcul le pourcentage d'échec des méthodes pour les différents points initiaux (tirage aléatoire de 4000 valeurs de T)



On voit que le nouveau point critique augmente le nombre d'échec avec la méthode de pénalisation de la fonction objectif. Une explication pour cela est que dans certains cas le point initial est dans le bon cas mais donne une valeur

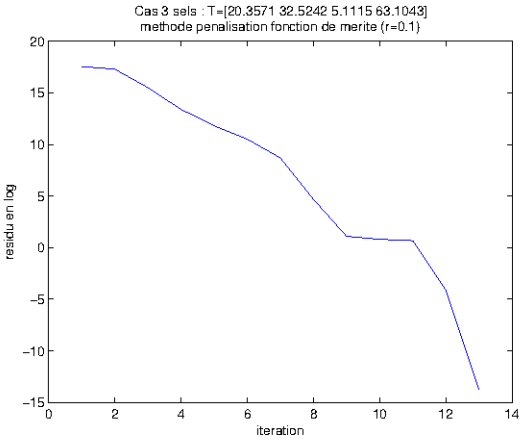
élevée du résidu et la méthode est donc tentée de changer de cas pour baisser le résidu et on ainsi on perd l'avantage que l'on avait. Une solution pour régler ce problème serait de réaliser quelques itérations en forçant la méthode à rester dans le cas fixé par la point initial.

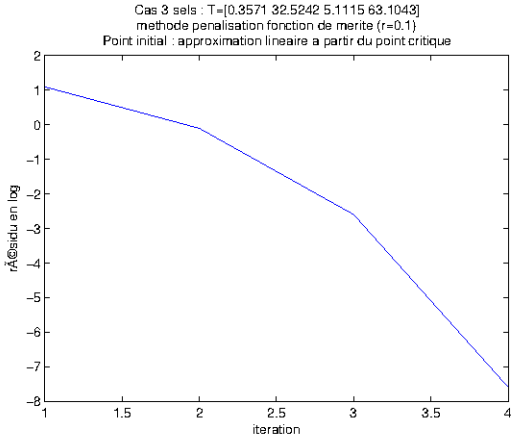
- Ensuite on compare le nombre d'itérations moyenne (un tirage aléatoire de 4000 valeurs de T) de la méthode de Newton.



On voit qu'avec le nouveau point initial on réalise 3 fois moins d'itérations.

- On compare la progression du cas 3 sels considéré précédemment avec le nouveau point initial





On voit que le phénomène de palier à été éliminé ce qui confirme le fait que l'on diminue le nombre d'itération à effectuer.

Part IV
Conclusion

Conclusion

Le travail de ce stage était centré sur l'étude du système de chimie utilisé notamment dans le contexte du transport réactif.

Nous avons vu dans la première partie comment choisir les variables du problème ainsi que les équations d'action de masse qui correspondent aux réactions de sorption et de précipitation-dissolution sous l'hypothèse que les réactions sont toujours à l'équilibre. C'est cette hypothèse qui permet au système de chimie d'avoir une structure polynomiale. Pour les réactions de précipitation-dissolution comme on a un phénomène de saturation le système contient un problème de complémentarité que l'on choisit de reformuler grâce à la C-fonction min.

Ensuite dans la seconde partie nous avons proposé une méthode analytique pour analyser le système de chimie qui ne concerne que les réactions de précipitation-dissolution. Cette méthode analytique possède l'avantage de donner des solutions exactes. La contrepartie de cette précision est que la méthode repose sur une approche combinatoire et sur la résolution de systèmes polynomiaux. L'étude du cas à 6 sels nous a permis de voir qu'à priori la combinatoire n'est pas un si gros problème de plus dans la pratique on a des raisons de penser qu'il n'y a pas beaucoup de minéraux intervenant dans la chimie. Il semble que le facteur limitant de cette approche est la résolution de systèmes polynomiaux ce qui nécessitera de nouveaux tests plus complexes.

Enfin dans la dernière partie sur l'étude numérique, nous avons vu comment reformuler et prendre en compte les difficultés spécifiques du système de chimie pour utiliser une méthode de Newton semi-lisse. En particulier nous avons sélectionné deux méthodes pour prendre en compte les contraintes de positivité sur c. La méthode à pas contraint bien que plus rapide semble échouer trop souvent pour être fiable. En échange la méthode à pénalisation de la fonction objectif est plus lente mais plus fiable. Nous avons pu remarquer que pour ces 2 méthodes il y avait un phénomène de palier qui ralenti la convergence. Ce phénomène de palier est causé par le fait que la méthode a besoin de temps pour se trouver dans le bon cas avant de converger vers la solution. La dernière section de cette partie a montré qu'en choisissant un nouveau point initial dans le bon cas on améliorait la convergence de la méthode à pénalisation de la fonction objectif. Ce nouveau point initial étant obtenu en utilisant le point critique calculé dans la partie analytique et en approximant les courbes critiques.

Ouverture

Nous avons pu au cours de ce travail observer plusieurs approches pour la résolution du système de chimie. Ces approches étaient concentrées sur la chimie où seules apparaissent des réactions de précipitation-dissolution, la question se pose donc des modifications à apporter si on généralise ce système en y ajoutant les réactions de sorption. Ces éventuelles modifications ont été rapidement discutées dans la section 12 et nécessiteront l'étude d'exemples pour estimer sa difficulté.

L'ajout de ces réactions de sorption ne devrait pas altérer les résultats obtenus dans l'étude numérique étant donné qu'elles ne présentent pas une difficulté supplémentaire.

On a vu en fin que l'étude analytique nous permettait d'obtenir un "bon" point initial. A nouveau il faudrait étudier de nouveaux exemples pour confirmer que l'heuristique utilisée pour déterminer le point initial est fiable.

Dans l'étude numérique nous avons observé que les méthodes étudiées avaient un pourcentage non-négligeable d'échecs. Dans ce travail nous avons choisi d'utiliser une méthode de line search par sa simplicité de mise en ?uvre mais utiliser à la place une méthode de région de confiance réputée plus robuste devrait permettre de réduire fortement ce taux d'échecs. Cela n'effacerait pas les travaux réalisés car les méthodes proposées pour prendre en compte les contraintes de positivité peuvent très bien s'adapter au cas de méthode de région de confiance.

Enfin cette étude sur le système de chimie pourrait être intégré dans un code de transport réactif pour être comparé à d'autres méthodes existantes.

Part V
Appendice

A Calculs des solutions : exemple 3 sels

Pour réaliser les calculs et obtenir les courbes critiques on se ramène à l'expression du système en fonction des variables c_i.

Cas 1 : 0 sel $=> p_1 = p_2 = p_3 = 0$ Le système devient :

$$\begin{cases} c_1 & = T_1 \\ c_2 & = T_2 \\ c_3 & = T_3 \\ c_1 c_4 & < K_1 \\ c_2 c_4 & < K_2 \\ c_2 c_3 c_4^3 & < K_3 \end{cases}$$

Les solutions sont $c_1 = T_1$, $c_2 = T_2$, $c_3 = T_3$ et $c_4 = T_4$, à condition de vérifier les inégalités :

$$\begin{cases} T_1 T_4 & < K_1 \\ T_2 T_4 & < K_2 \\ T_2 T_3 T_4^3 & < K_3 \end{cases}$$

En utilisant la neutralité électrique, on a :

$$\begin{cases} T_1^2 + T_1(T_2 + 2T_3) - K_1 & < 0 \\ T_2^2 + T_2(T_1 + 2T_3) - K_2 & < 0 \\ T_2 T_3 (T_1 + T_2 + 2T_3)^3 - K_3 & < 0 \end{cases}$$

Remarque 2 : Soit un polynôme de la forme $x^2 + x(T_i + 2T_3) - K_j$, ce polynôme possède deux racines $x_{ij} = -\frac{1}{2}T_i - T_3 \underset{-}{+} \frac{1}{2}\sqrt{T_i^2 + 4T_iT_3 + 4T_3^2 + 4K_j}$, le polynôme est négatif entre les racines. Si on considère des valeurs de T positives, alors il y a une seule racine positive : $x_{ij} = -\frac{1}{2}T_i - T_3 + \frac{1}{2}\sqrt{T_i^2 + 4T_iT_3 + 4T_3^2 + 4K_j}$.

Remarque 3 : On a des formules pour calculer les racines du polynôme $T_2T_3(T_1 + T_2 + 2T_3)^3 - K_3 = 0$ mais la solution est compliquée.

Cas 2 : 3 sels $=> p_1 > 0$, $p_2 > 0$, $p_3 > 0$ Le système devient :

$$\begin{cases} c_1 & < T_1 \\ c_2 & < T_2 - T_3 + c_3 \\ c_3 & < T_3 \\ c_1 c_4 & = K_1 \\ c_2 c_4 & = K_2 \\ c_2 c_3 c_4^3 & = K_3 \end{cases}$$

On utilise l'expression $c_4 = c_1 + c_2 + 2c_3$ pour obtenir un polynôme de degré 3 en c_4:

$c_4^3 - c_4(K_1 + K_2) - 2\frac{K_3}{K_2}$, or il existe une unique solution positive que l'on a déjà vue $c_4 = T_4^p$. Les solutions du système sont $c_1 = T_1^p$, $c_2 = T_2^p$, $c_3 = T_3^p$ et $c_4 = T_4^p$, sous les contraintes :

$$\begin{cases} T_1^p & < T_1 \\ T_2^p - T_3^p + T_3 & < T_2 \\ T_3^p & < T_3 \end{cases}$$

Cas 3 : 1 sel (p1) $=> p_1 > 0$**,** $p_2 = 0$**,** $p_3 = 0$ Le système devient :

$$\begin{cases} c_1 & < T_1 \\ c_2 & = T_2 \\ c_3 & = T_3 \\ c_1 c_4 & = K_1 \\ c_2 c_4 & < K_2 \\ c_2 c_3 c_4^3 & < K_3 \end{cases}$$

On utilise l'expression $c_4 = c_1 + c_2 + 2c_3$, on obtient un polynôme de degré 2 en c_1; $c_1^2 + c_1(c_2 + 2c_3) - K_1 = c_1^2 + c_1(T_2 + 2T_3) - K_1$, il existe une unique racine positive $\bar{c_1}$ pour T positif que l'on a rencontré dans le cas 1 et qui dépend de T_2 et T_3. Les solutions de ce système sont $c_1 = \bar{c_1}$, $c_2 = T_2$, $c_3 = T_3$et $c_4 = \bar{c_1} + T_2 + 2T_3$, à condition que :

$$\begin{cases} \bar{c_1} & < T_1 \\ \bar{c_1} c_4 & = K_1 \\ T_2 c_4 & < K_2 \\ T_2 T_3 c_4^3 & < K_3 \end{cases}$$

On peut éliminer c_4, on obtient :

$$\begin{cases} \bar{c_1} & < T_1 \\ \frac{K_2}{K_1}\bar{c_1} - T_2 & > 0 \\ \frac{K_3}{T_2 K_1^3}\bar{c_1}^3 - T_3 & > 0 \end{cases}$$

On peut noter que $\frac{K_2}{T_3+\sqrt{T_3^2+K_1+K_2}}$ est l'unique racine réelle positive de $\frac{K_2}{K_1}\bar{c_1} - T_2$

Cas 4 : 1 sel (p2) $=> p_1 = 0,\ p_2 > 0,\ p_3 = 0$ Le système devient :

$$\begin{cases} c_1 & = T_1 \\ c_2 & < T_2 \\ c_3 & = T_3 \\ c_1c_4 & < K_1 \\ c_2c_4 & = K_2 \\ c_2c_3c_4^3 & < K_3 \end{cases}$$

On utilise l'expression $c_4 = c_1 + c_2 + 2c_3$, on obtient un polynôme de degré 2 en c_2; $c_2^2 + c_2(c_1 + 2c_3) - K_2 = c_2^2 + c_2(T_1 + 2T_3) - K_2$, il existe une unique racine positive $\bar{c_2}$ pour T positif que l'on a rencontré dans le cas 1 et qui dépend de T_1 et T_3. Les solutions de ce système sont $c_2 = \bar{c_2}$, $c_1 = T_1$, $c_3 = T_3$et $c_4 = \bar{c_2} + T_1 + 2T_3$, à condition que :

$$\begin{cases} \bar{c_2} & < T_2 \\ \bar{c_2}c_4 & = K_2 \\ T_1c_4 & < K_1 \\ \bar{c_2}T_3c_4^3 & < K_3 \end{cases}$$

On peut éliminer c_4, on obtient :

$$\begin{cases} T_1 - \frac{K_1}{K_2}\bar{c_2} & < 0 \\ \bar{c_2} & < T_2 \\ T_3 - \frac{K_3}{K_2^3}\bar{c_2}^2 & < 0 \end{cases}$$

On peut noter que $\frac{K_1}{T_3+\sqrt{T_3^2+K_1+K_2}}$ est l'unique racine réelle positive de $\frac{K_1}{K_2}\bar{c_2} - T_1$

Cas 5 : 1 sel (p3) $=> p_1 = 0,\ p_2 = 0,\ p_3 > 0$ Le système devient :

$$\begin{cases} c_1 & = T_1 \\ c_2 & = T_2 - T_3 + c_3 \\ c_3 & < T_3 \\ c_1c_4 & < K_1 \\ c_2c_4 & < K_2 \\ c_2c_3c_4^3 & = K_3 \end{cases}$$

On obtient un polynôme de degré 5 en c_3: $K_3 - (T_2 - T_3 + c_3)c_3(T_1 + T_2 - T_3 + 3c_3)^3$. Soit la solution du polynôme de degré 5 en c_3: $f(T_1, T_2, T_3)$. Alors les solutions du systèmes sont $c_3 = f(T_1, T_2, T_3)$, $c_2 = T_2 - T_3 + f(T_1, T_2, T_3)$ et $c_1 = T_1$, sous les conditions :

$$\begin{cases} f(T_1,T_2,T_3) & < T_3 \\ T_1(T_1+T_2-T_3+3f(T_1,T_2,T_3))-K_1 & < 0 \\ (T_2-T_3+f(T_1,T_2,T_3))(T_1+T_2-T_3+3f(T_1,T_2,T_3))-K_2 & < 0 \end{cases}$$

Cas 6 : 2 sels (p1) et (p2) $=> p_1 > 0$, $p_2 > 0$, $p_3 = 0$ Le système devient :

$$\begin{cases} c_1 & < T_1 \\ c_2 & < T_2 \\ c_3 & = T_3 \\ c_1c_4 & = K_1 \\ c_2c_4 & = K_2 \\ c_2c_3c_4^3 & < K_3 \end{cases}$$

On utilise l'expression $c_4 = c_1 + c_2 + 2c_3$, on obtient un polynôme de degré 2 en c_4 : $c_4^2 - 2c_4c_3 - K_1 - K_2$, il existe une unique racine positive $c_4 = c_3 + \sqrt{c_3^2 + K_1 + K_2}$. Les solutions du système sont $c_3 = T_3$, $c_4 = T_3 + \sqrt{T_3^2 + K_1 + K_2}$, $c_1 = \frac{K_1}{c_4}$ et $c_2 = \frac{K_2}{c_4}$, sous les contraintes :

$$\begin{cases} c_1 & < T_1 \\ c_2 & < T_2 \\ K_2c_3c_4^2 & < K_3(*) \end{cases}$$

On injecte la solution $c_4 = T_3 + \sqrt{T_3^2 + K_1 + K_2}$. dans (*), on obtient :

$K_3 - K_2T_3(T_3 + \sqrt{T_3^2 + K_1 + K_2})^2 > 0$, on remarque qu'il existe une unique racine réelle au polynôme $K_3 - K_2T_3(\ T_3 + \sqrt{T_3^2 + K_1 + K_2})^2$ et de plus cette racine est T_3^p. Les contraintes deviennent :

$$\begin{cases} c_1 & < T_1 \\ c_2 & < T_2 \\ T_3^p & > T_3 \end{cases}$$

Cas 7 : 2 sels (p1) et (p3) $=> p_1 > 0$, $p_3 > 0$, $p_2 = 0$ Le système devient :

$$\begin{cases} c_1 & < T_1 \\ c_2 & = T_2 - T_3 + c_3 \\ c_3 & < T_3 \\ c_1c_4 & = K_1 \\ c_2c_4 & < K_2 \\ c_2c_3c_4^3 & = K_3 \end{cases}$$

On utilise l'expression $c_4 = c_1 + c_2 + 2c_3 = \frac{K_1}{c_4} + T_2 - T_3 + 3c_3$, on obtient un polynôme de degré 2 en c_4: $c_4^2 + c_4(T_3 - T_2 - 3c_3) - K_1$ qui possède une unique racine réelle positive, $\bar{c_4} = \frac{1}{2}(-T_3 + T_2 + 3c_3 + \sqrt{9c_3^2 + 6c_3T_2 - 6c_3T_3 + T_2^2 - 2T_2T_3 + T_3^2 + 4K_1})$.

On utilise cette expression de c_4 en fonction de $c_3$pour obtenir un polynôme de degré 5 en c_3: $(T_2 - T_3 + c_3)c_3(\bar{c_4})^3 - K_3$. On suppose qu'il existe une unique racine réelle positive à ce polynôme et on note $c_3 = g(T_2, T_3)$. Les solutions du système sont : $c_3 = g(T_2, T_3)$, $c_4 = c_4^*(g(T_2, T_3))$,$c_2 = T_2 - T_3 + g(T_2, T_3)$ et $c_1 = \frac{K_1}{c_4}$, sous les conditions :

$$\begin{cases} c_1 & < T_1 \\ c_3 & < T_3 \\ c_2c_4 & < K_2 \end{cases}$$

Cas 8 : 2 sels (p2) et (p3) $=> p_1 = 0$, $p_2 > 0$, $p_3 > 0$ Le système devient :

$$\begin{cases} c_1 & = T_1 \\ c_2 & < T_2 - T_3 + c_3 \\ c_3 & < T_3 \\ c_1c_4 & < K_1 \\ c_2c_4 & = K_2 \\ c_2c_3c_4^3 & = K_3 \end{cases}$$

On utilise l'expression $c_4 = c_1 + c_2 + 2c_3 = T_1 + \frac{K_2}{c_4} + 2\frac{K_3}{K_2c_4^2}$, on obtient un polynôme de degré 3 en c_4: $c_4^3 - c_4^2T_1 - c_4K_2 + 2\frac{K_3}{K_2}$, il existe une unique racine réelle positive à ce polynôme que l'on nomme $c_4^*(T_1)$. Les solutions du système sont $c_4 = c_4^*(T_1)$, $c_1 = T_1$,$c_2 = \frac{K_2}{c_4^*}$ et $c_3 = \frac{K_3}{K_2(c_4^*)^2}$, sous les conditions :

$$\begin{cases} c_2 & < T_2 - T_3 + c_3 \\ c_3 & < T_3 \\ c_1c_4 & < K_1 \end{cases}$$

B Benchmark MoMas Transport Réactif "Easy Test Case"

On peut notamment se servir comme exemple du cas test "Easy test case" proposé par le groupe de recherche MoMas dans le contexte du transport réactif [1], qui propose plusieurs difficultés d'une part pour la grande différence de magnitude entre les constantes et d'autre part les coefficients de la matrices stœchiométrique sont importants également.

Table 4 : Equilibrium for easy test case

	X_1	X_2	X_3	X_4	S	K
C_1	0	-1	0	0	0	1.00E-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	1.00E+35
CS_1	0	3	1	0	1	1.00E+6
CS_2	0	-3	0	1	2	1.00E-01
Total ($m.L^{-3}$)	T_1	T_2	T_3	T_4	TS	
Initial for medium A	0	-2	0	2	1	
Initial for medium B	0	-2	0	2	10	
Injection $t \in [0, 5000]$	Imposed total concentration at inflow boundary					
Inflow for 1D	0.3	0.3	0.3	0		
Zone 1 for 2D	0.3	0.3	0.3	0		
Zone 2 for 2D	0.3	0.3	0.3	0		
Leaching $t \in [5000, ...]$	Imposed total concentration at inflow boundary					
Inflow for 1D	0	-2	0	2		
Zone 1 for 2D	0	-2	0	2		
Zone 2 for 2D	0	-2	0	2		

Qui nous donne un système de la forme :

$$\begin{cases} \alpha_1 = K_1 c_2^{-1} \\ \alpha_2 = K_2 c_2 c_3 \\ \alpha_3 = K_3 c_2^{-1} c_4 \\ \alpha_4 = K_4 c_2^{-4} c_3 c_4 \\ \alpha_5 = K_5 c_2^4 c_3^3 c_4 \\ \beta_1 = K_6 c_2^3 c_3 s \\ \beta_2 = K_7 c_2^{-3} c_4 s^2 \\ T_1 = c_1 \\ T_2 = c_2 - \alpha_1 + \alpha_2 - \alpha_3 - 4\alpha_4 + 4\alpha_5 + 3\beta_1 - 3\beta_2 \\ T_3 = c_3 + \alpha_2 + \alpha_4 + 3\alpha_5 + \beta_1 \\ T_4 = c_4 + \alpha_3 + 3\alpha_4 + \alpha_5 + \beta_2 \\ T_5 = s + \beta_1 + 2\beta_2 \end{cases}$$

L'approche formel de ce calcul par des bases de Gröbner a été mentionné

dans [17].

On remplace les α et les β dans les lois de conservations. L'équation $T_1 = c_1$ étant trivial on peut la mettre de côté. On obtient un système polynomial de quatre équations à 4 inconnues (c_2, c_3, c_4 et s).

$$\begin{cases} -4K_5c_3^3c_4c2^8 - 3K_6c_3sc_2^7 + (-c_3K_2 - 1)c_2^5 + T_2c_2^4 + (c_4K_2 + K_1)c_2^3 + 3K_7c_4s^2c_2 + 4K_4c_3c_4^3 & = 0 \\ -3K_5c_3^3c_4c_2^8 - K_6c_3sc_2^7 - K_2c_3c_2^5 + (T_3 - c_3) * c_2^4 - K_4c_3c_4^3 & = 0 \\ -K_5c_3^3c_4c_2^8 + (T_4 - c_4)c_2^4 - K_2c_4c_2^3 - K_7c_4s^2c_2 - 3K_4c_3c_4^3 & = 0 \\ -K_6c_3sc_2^6 + (T_s - s)c_2^3 - 2K_7c_4s^2 & = 0 \end{cases}$$

References

[1] J. Carrayrou A. Dimier C.J. Van Duijn M. Kern P. Knabner N. Leterrier A. Bourgeat, S. Bryant. Gdr momas, benchmark, reactive transport. 2008.

[2] Radu Serban Aaron M. Collier, Alan C. Hindmarsh and Carol S. *User Documentation for kinsol v2.7.0.* Woodward Center for Applied Scientific Computing Lawrence Livermore National Laboratory, 2011.

[3] C.M. Bethke. *Geochemical Reaction Modeling : Concepts and Applications.* Oxford University Press US, 1996.

[4] H. Buchholzer. *The Semismooth Newton Method for the Solution of Reactive Transport Problems Including Mineral Precipitation-Dissolution Reactions.* PhD thesis, University of Würzburg, 2011.

[5] J. Erhel C. de Dieuleveult. A global approach to reactive transport: application to the momas benchmark. *Computational Geoscience*, 14:451–464, 2010.

[6] M. Kern C. de Dieuleveult, J. Erhel. A global strategy for solving reactive transport equations. *Journal of Computational Physics*, 228, 2009.

[7] S.C. Smith C. Liu, J.M. Zachara. A cation exchange model to describe cs+ sorption at high ionic strength in subsurface sediments at hanford site. *Journal of Contaminant Hydrology*, 2004.

[8] G.D. Bennett C. Zheng. *Applied Contaminant Transport Modeling - second edition.* John Wiley and Sons, 2002.

[9] D. Postma. C.A.J. Appelo. *Geochemistry, groundwater and pollution, second edition.* A. A. Balkema, 2005.

[10] K.T.B. MacQuarrie C.I. Steefel. Approaches to modeling of reactive transport in porous media. *Reviews in Mineralogy and Geochemistry*, 1996.

[11] C.T.Kelley. *Iterative Methods for Linear and Nonlinear Equations.* SIAM, 1995.

[12] C.T.Kelley. *Solving Nonlinear Equations with Newton's Method.* SIAM, 2003.

[13] F.M.M. Morel. D.A. Dzombak. *Surface Complexation Modeling: Hydrous Ferric Oxide.* Wiley-Interscience, 1990.

[14] C. de Dieuleveult. *Un modèle numérique global et performant pour le couplage géochimie-transport.* PhD thesis, Université de Rennes 1, 2008.

[15] J.G. Hering F. Morel. *Principles and application of aquatic chemistry.* Wiley- IEEE, 1993.

[16] P.C. Lichtner G.E. Hammond, A.J. Valocchi. Application of jacobian-free newton-krylov with physics-based preconditioning to biogeochemical transport. *Advances in Water Resources*, 2005.

[17] J. Houot M. Kern A. Semin J.-B. Apoung, P. Have. Reactive transport in porous media. pages Vol.28 p.227–245. ESAIM: PROCEEDINGS, August 2009.

[18] P. Knabner J. Hoffmann, S. Kräutle. A general reduction scheme for reactive transport in porous media. *Computational Geoscience*, 2012.

[19] B. Bussiere J. Molson, M. Aubertin. Reactive transport modelling of acid mine drainage within discretely fractured porous media: Plume evolution from a surface source zone. *Environmental Modelling & Software 38*, 2012.

[20] V. Lagneau P.P. Goblet J. van der Lee, L.L. De Windt. Moduleoriented modeling of reactive transport with hytec. *Computers and Geosciences*, 2003.

[21] S. Gordon JF.J. Zeleznik. Calculation of complex chemical equilibria. *Industrial and Engineering Chemistry*, 1968.

[22] C. Kanzow. An active-set newton method for constrained nonlinear systems. 1999.

[23] C. Kanzow. Strictly feasible equation-based methods for mixed complementarity problems. 1999.

[24] S. Kräutle. *General Multi-Species Reactive Transport Problems in Porous Media : Efficient Numerical Approaches and Existence of Global Solutions.* Habilitation thesis, University orf Erlangen-Nuremberg, 2008.

[25] D.W. Blowes K.U. Mayer, E.O. Frind. Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. *Water Resources Research*, 2002.

[26] P.C. Lichtner. Continuum formulation of multicomponent-multiphase reactive transport. *Reviews in Mineralogy and Geochemistry*, 1996.

[27] J. Rubin. Transport of reacting solutes in porous media ; relation between mathematical nature of problem formulation and chemical nature of reactions. *Water Ressources Research*, 1983.

[28] P. Knabner S. Kräutle. A reduction scheme for coupled multicomponent transport-reaction problems in porous media : generalization to problems with heterogeneous equilibrium reactions. *Water Resources Research*, 2007.

[29] C. Ayora M.W. Saaltink S. Molins, J. Carrera. A formulation for decoupling components in reactive transport problems. *Water Resources Research*, 2004.

[30] S. Sabit. *Les méthodes numériques de transport réactif.* PhD thesis, Université Rennes 1, 2014.

[31] M. Ulbrich. *Semismooth Newton Methods for variational inequalities and constrained optimization problems in function spaces.* MOS-SIAM Series on Optimization, 2011.