Calcul Scientifique pour l'Environnement
Cours ENSTA

Bruno Sportisse et Vivien Mallet [1]

25 février 2005

[1]Centre d'Enseignement et de Recherche sur l'Environnement Atmosphérique. Laboratoire Commun EDF R&D/ENPC. sportiss@cerea.enpc.fr

Ces notes de cours correspondent à un enseignement dispensé à l'Ecole Nationale Supérieure des Techniques Avancées (ENSTA), *Calcul Scientifique pour l'Environnement* pour l'année 2005.

Nous remercions Laurent Mortier pour avoir proposé l'organisation de ce cours.

Nous tenons à remercier Karine Béranger pour les discussions autour des petites classes associées.

Bruno Sportisse et Vivien Mallet

Introduction

Modélisation et simulation en environnement/géophysique

Les problèmes rencontrés en modélisation de l'environnement, notamment géophysique (océanographie, hydrologie, météorologie, pollution) ont un certain nombre de spécificités :

1. ils sont d'abord très fortement *multi-disciplinaires* couplant mécanique des fluides (en clair les équations de Navier-Stokes pour décrire l'écoulement du fluide considéré : air ou eau), physique (pour décrire le comportement microphysique, par exemple des particules), chimie ou biologie (pour l'évolution des espèces considérées), transfert radiatif (les rayonnements terrestre et solaire), etc.

   Un point clé est alors la nécessité d'effectuer des *couplages de modèles*, qui constituent le point de développement de ces domaines.

2. la problématique des *rétroactions* (des "feedbacks") est dans ce contexte cruciale, la question étant de préciser jusqu'à quel point il est pertinent de coupler.

   Un exemple caractéristique est fourni par l'interaction chimie/rayonnement/nuage (et le rôle des aérosols dans l'atmosphère) qui est l'un des points encore largement ouverts pour l'évaluation du changement climatique.

3. les problèmes traités sont souvent de *très grande dimension* (des centaines d'espèces chimiques pour la chimie atmosphérique).

4. les problèmes sont *multi-échelles*, de nombreuses échelles étant à considérer simultanément : par exemple, en chimie atmosphérique, les échelles temporelles liées aux processus chimiques s'étendent sur plusieurs ordres de grandeurs (des espèces radicalaires aux espèces stables), les échelles spatiales de quelques nanomètres (la formation des aérosols) à l'échelle de l'écoulement géophysique.

   Ceci conduit à de nombreuses difficultés d'ordre numérique.

5. de manière induite, la problématique de la *paramétrisation* des processus est un point clé : comment représenter des processus définis à petite échelle (en temps et en espace) dans des modèles dont le "grain" (spatial et temporel) est de fait relativement grossier ?

6. les milieux représentés sont très hétérogènes et de très grandes incertitudes existent dans les données nécessaires à l'utilisation des modèles numériques (conditions initiales de problèmes d'évolution, conditions aux limites, description des milieux géophysiques : topographie, occupation du terrain, etc).

   Dans ce contexte, le couplage entre modèles numériques et données observées, fournies par des réseaux de mesures (terrestres mais aussi de plus en plus satellitaires - le domaine de l'*Observation de la Terre*-) est une approche incontournable. On parle alors d'*assimilation de données*, pour laquelle des approches méthologiques sont nécessaires. Ce point est essentiel notamment pour des applications des modèles à la *prévision*.

De manière schématique, la modélisation dans ce domaine s'articule autour des trois activités suivantes :

1. la modélisation physique à proprement parler, pour laquelle la problématique de la paramétrisation sous-maille et la confrontation aux données de terrain sont cruciales ;

2. la simulation numérique des modèles construits ;
3. l'assimilation de données.

Ce cours s'insère dans un ensemble de trois cours et a pour objet le second thème. La modélisation, avec l'exemple spécifique de la pollution atmosphérique, fait l'objet du cours [47], l'assimilation de données du cours [46].

Dans un premier temps, le cours s'est focalisé sur les *modèles de dispersion réactive de traceurs* (par exemple dans l'atmosphère ou dans un autre milieu). Cet exemple est spécifique mais permet de balayer un grand nombre de méthodes largement génériques et utilisées dans d'autres applications. L'objectif de ce cours est, de manière plus générale, de donner les principaux éléments de calcul scientifique appliqués à la simulation numérique des problèmes que l'on rencontre en environnement géophysique.

Organisation

Ce cours est organisé de la manière suivante.

Dans le chapitre 1, on présente (rappelle ?) le modèle de dispersion réactive.

Dans le chapitre 2, on étudie une classe de méthodes couramment utilisées dans ce domaine, les méthodes de *séparation d'opérateurs* (splitting).

Dans le chapitre 3, le traitement spécifique des termes réactifs (chimiques par exemple) est abordé, notamment autour de la problématique des modèles *raides* (présentant une grande dispersion des échelles de temps).

Enfin, la résolution des termes de transport (advection et diffusion) est traitée dans le chapitre 4.

Table des matières

Chapitre 1

Equation de dispersion réactive

L'objectif de ce chapitre est de rappeler brièvement les équations de dispersion réactive de traceurs dans un milieu géophysique (air, eau).

On rappelle l'équation de dispersion réactive dans la section 1 en détaillant les processus d'advection, de diffusion et de réactions. Un point clé est l'hypothèse de dilution qui consiste à découpler les équations de la dynamique du fluide de celles de l'évolution des traceurs.

Les principaux processus sont ensuite classifiés dans la section 2 (c'est ce qui les rattachera à des familles d'algorithmes numériques).

Enfin, on conclut en discutant brièvement du choix des méthodes de discrétisation spatiale.

1.1 Equations de dispersion réactive

1.1.1 Hypothèse de dilution

On cherche à décrire dans un milieu donné (atmosphère, océan, fleuve) la dispersion d'un jeu d'espèces chimiques (ou biologiques), supposées réagir entre elles.

En toute rigueur, l'évolution du système couplé (fluide+traceurs) est donnée par les équations de Navier-Stokes réactives. On fait néanmoins de manière classique une *hypothèse de dilution* qui consiste à découpler d'une part la dynamique du fluide, d'autre part les concentrations de traceurs. Ceci revient notamment à négliger dans l'équation d'évolution d'énergie interne (ou de température) la contribution dûe aux réactions chimiques et à figer l'interaction matière/rayonnement. Par exemple, dans le cas de l'atmosphère, la première approximation est bien vérifiée dans la troposphère alors que la seconde revient à négliger un moteur clé de la dynamique atmosphérique.

Dans ce cas de figure, les champs dynamiques (vent, diffusion, température, humidité de l'air pour le cas atmosphérique) sont donc calculés indépendamment ou paramétrisés, et sont utilisés comme des données connues dans l'équation de dispersion pour les traceurs considérés.

Il est à noter pour finir qu'une telle hypothèse n'est évidemment plus valable pour des systèmes où le couplage chimie/dynamique est beaucoup plus important (par exemple en

combustion). Les équations de Navier-Stokes réactives doivent alors être traitées, ce qui dépasse le cadre de ce cours et fait appel pour partie à d'autres méthodes.

1.1.2 Equations d'advection-diffusion-réaction

Dans le cadre d'une hypothèse de dilution, l'évolution des traceurs indicés par i obéit alors à un système d'Equations aux Dérivées Partielles donné par :

$$\frac{\partial c_i}{\partial t} + \mathrm{div}(V(x,t)c_i) = \mathrm{div}(K_{molec}\nabla c_i) + \chi_i(c, T(x,t), t) + S_i(x,t) \tag{1.1}$$

où x et t désignent respectivement les coordonnées d'espace et de temps, c est le vecteur des concentrations d'espèces (indicées par i), $V(x,t)$ est le champ de vitesse du fluide, K_{molec} est la matrice de diffusion moléculaire (a priori non diagonale, du fait de la diffusion inter-moléculaire), $T(x,t)$ est le champ de température.

$S_i(x,t)$ est le terme source pour l'espèce i, qui modélise le cas échéant l'émission par source fixe. Dans le cas atmosphérique (figure 1.1), ceci correspond typiquement à des émissions par cheminées d'usine ; dans le cas hydrologique, à une source ponctuelle de pollution.

Enfin, χ_i désigne le taux de production chimique de l'espèce i, sur lequel on reviendra plus spécifiquement par la suite.

Pour terminer, l'équation précédente n'est en réalité valable que dans le cas d'un fluide incompressible (densité ρ constante). Dans le cas général, la densité du fluide porteur vérifie l'équation de continuité :

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho V) = 0 \tag{1.2}$$

et la "bonne" variable pour le traceur est son rapport de mélange que l'on notera $m_i = c_i/\rho$, dont l'évolution est donnée par :

$$\frac{\partial m_i}{\partial t} + V \cdot \nabla m_i = \frac{1}{\rho}\mathrm{div}(K_{molec}\nabla(\rho m_i)) + \frac{\chi_i(\rho m, T(x,t), t) + S_i(x,t)}{\rho} \tag{1.3}$$

1.1.3 Modèles moyens

En réalité, cette équation d'évolution, si elle est valide au niveau "microscopique", n'est pas applicable telle quelle pour des écoulements turbulents. Ceux-ci peuvent notamment être caractérisés par une grande disparité des échelles spatiales : par exemple, pour la cas d'une turbulence d'origine dynamique (cisaillement de vitesse -typiquement le cas d'une couche limite dynamique), l'analyse de Kolmogorov donne un rapport d'échelle entre la plus petite échelle caractéristique (l) et la plus grande (L) en fonction du nombre de Reynolds (Re) selon :

$$\frac{L}{l} \simeq Re^{3/4} \tag{1.4}$$

Pour des écoulements fortement turbulents ($Re \gg 1$), il est évidemment impossible de simuler l'ensemble des échelles en 3 dimensions. On a alors recours de manière classique à des approches de moyennisation.

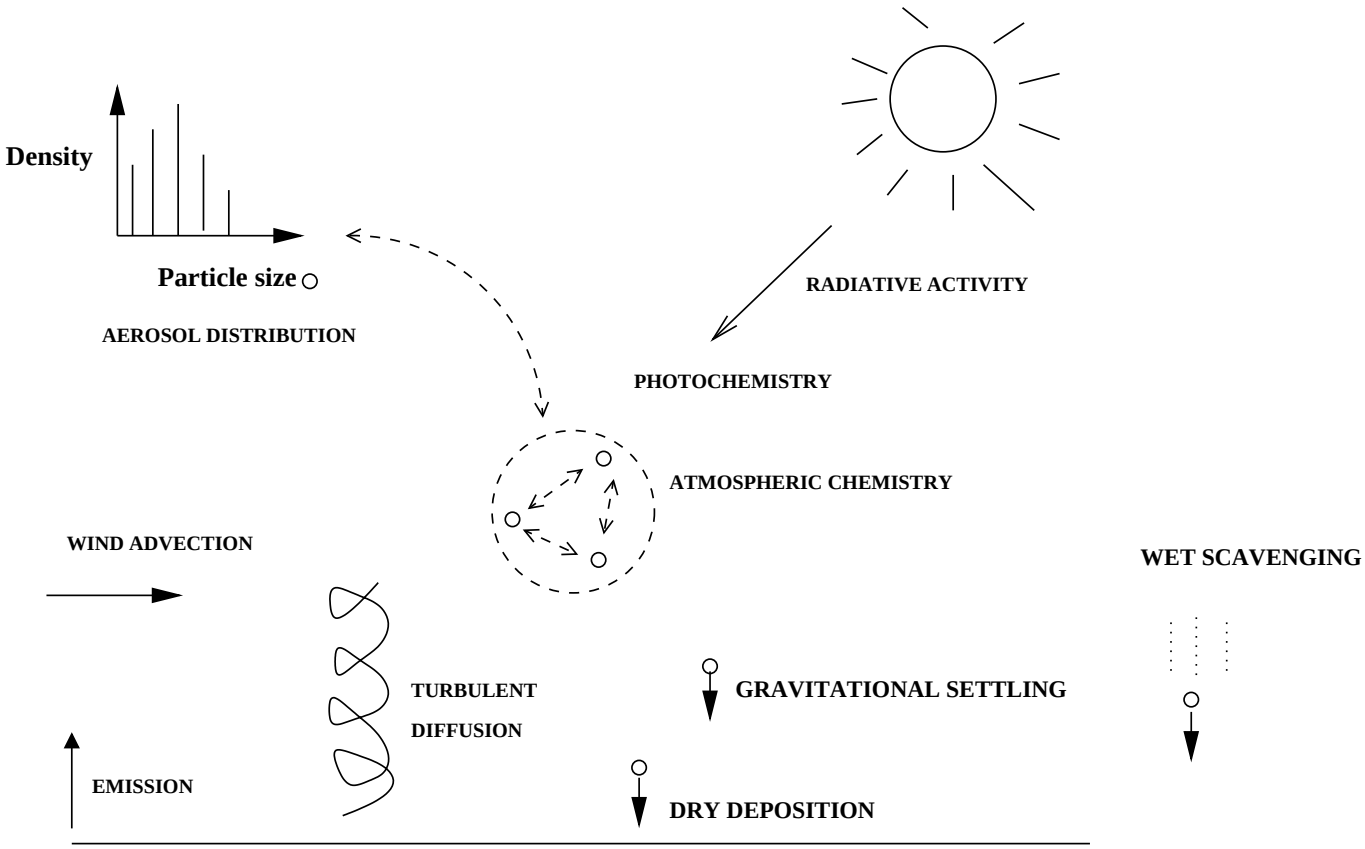

FIG. 1.1 – Processus décrits dans un modèle de Chimie-Transport.

En omettant de préciser la définition rigoureuse de l'opérateur de moyenne (en un sens spatial, en un sens ergodique, avec ou sans une pondération par la densité -moyenne de Favre-...), on suppose donc à présent que les champs étudiés se décomposent selon :

$$\Psi = \langle \Psi \rangle + \Psi^{'} \tag{1.5}$$

avec $\langle \Psi \rangle$ une grandeur moyenne et $\Psi^{'}$ une fluctuation. Parmi les principales propriétés "demandées" à l'opérateur de moyenne, on retiendra qu'il commute avec les opérateurs de dérivation (en temps et en espace) et que $\left\langle \Psi^{'} \right\rangle = 0$.

De manière directe, une telle décomposition appliquée à l'équation précédente pour les traceurs c et pour le champ de vitesse V conduit après moyennisation à l'équation :

$$\frac{\partial \langle c_i \rangle}{\partial t} + \mathrm{div}(\langle V(x,t) \rangle \langle c_i \rangle) = \mathrm{div}(K_{molec} \nabla \langle c_i \rangle) + \langle \chi_i(c, T(x,t), t) \rangle + \langle S_i(x,t) \rangle - \mathrm{div}(\left\langle c_i^{'} V^{'} \right\rangle) \tag{1.6}$$

Il est direct de remarquer que les termes linéaires sont transposés tels quels dans l'équation moyennée. Les termes non-linéaires (en l'occurrence quadratiques) font apparaître des corrélations entre variables (la moyenne de produits de fluctuations). Le problème de la *fermeture* des équations moyennées revient alors à exprimer ces corrélations en fonction des grandeurs résolues (les valeurs moyennes).

Revenons sur les deux principaux termes à fermer dans l'équation précédente :

1. La moyennisation de l'équation de continuité pour l'espèce i conduit à l'introduction d'un terme de flux turbulent, non spécifié, associé au terme d'advection : $\mathrm{div} \left\langle c_i^{'} V^{'} \right\rangle$.
   Le problème de la fermeture des équations est résolu de manière classique à l'aide de la théorie du gradient (ou théorie K) qui revient à exprimer le flux turbulent d'une quantité advectée comme inversement proportionnel au gradient de la valeur moyennée. Pour un champ Ψ, la paramétrisation est donc du type :

   $$\left\langle \Psi^{'} V^{'} \right\rangle = -K_{turb}^{\Psi}(x,t) \nabla \langle \Psi \rangle \tag{1.7}$$

   avec K_{turb}^{Ψ} la diffusion turbulente dépendant de l'espace et du temps (en pratique donnée en fonction des champs dynamiques et de leurs gradients).
   Il est à noter que cette paramétrisation appliquée à la concentration c_i ou à la fraction massique m_i ne conduit pas à la même fermeture. Comme l'équation de continuité moyennée s'écrit usuellement sous la forme :

   $$\frac{\partial \langle \rho \rangle}{\partial t} + \mathrm{div}(\langle \rho \rangle \langle V \rangle) = 0 \tag{1.8}$$

   il est plus cohérent d'appliquer la paramétrisation à la fraction massique selon :

   $$\left\langle m_i^{'} V^{'} \right\rangle = -K_{turb} \nabla \left\langle m_i^{'} \right\rangle \tag{1.9}$$

   Avec $c_i = \rho m_i$, on fait l'approximation usuelle :

   $$c_i^{'} V^{'} = \langle \rho \rangle m_i^{'} V^{'} + \langle m_i \rangle \rho_i^{'} V^{'} \simeq \langle \rho \rangle m_i^{'} V^{'} \tag{1.10}$$

La paramétrisation donne alors pour la concentration :

$$\left\langle c_i^{'} V^{'} \right\rangle = - \langle \rho \rangle K_{turb} \nabla \frac{\left\langle c_i^{'} \right\rangle}{\langle \rho \rangle} \tag{1.11}$$

Pour le fluide (l'air), supposé non réactif, l'équation d'advection moyennée redonne l'équation de continuité (1.8) avec $c_i = \rho$, ce qui n'aurait pas été le cas en appliquant directement la paramétrisation à la concentration. On aurait alors fait apparaître un membre de droite dans l'équation de continuité moyennée.

Notons que la diffusion turbulente est supposée être la même pour toutes les espèces, la diffusion inter-espèces n'étant pas prise en compte ; autrement dit, la matrice K_{turb} est diagonale.

En pratique, on fait en général l'approximation $K_{turb} \gg K_{molec}$. Par exemple, pour le cas atmosphérique, la diffusion est uniquement turbulente en dehors d'une couche laminaire à proximité du sol, ce qui justifie cette simplification.

2. Le processus de moyennisation conduit de même, en toute rigueur, à un problème de fermeture pour la chimie non linéaire.

   A une réaction bimoléculaire de réactants notés symboliquement X_i et X_j est associée une production chimique proportionnelle à (voir annexe) :

$$\langle c_i c_j \rangle = \langle c_i \rangle \langle c_j \rangle + \left\langle c_i^{'} c_j^{'} \right\rangle \tag{1.12}$$

   où le terme de corrélation $\left\langle c_i^{'} c_j^{'} \right\rangle$ est inconnu.

   Ces termes sont habituellement négligés et on fait donc dans l'équation de dispersion l'approximation (dite parfois du *réacteur homogène* -well stirred tank reactor-) :

$$\langle \chi(c) \rangle \simeq \chi(\langle c \rangle) \tag{1.13}$$

   Une condition de validité est typiquement que les temps caractéristiques de la chimie sont beaucoup plus grands que ceux associés aux processus d'homogénéisation. Notons que le terme de corrélation négligé correspond à ce que l'on appelle un terme de *ségrégation* :

$$\langle c_i c_j \rangle = \langle c_i \rangle \langle c_j \rangle (1 + I_s) \ , \quad I_s = \frac{\left\langle c_i^{'} c_j^{'} \right\rangle}{\langle c_i \rangle \langle c_j \rangle} \tag{1.14}$$

   avec I_s l'*intensité de ségrégation* (dont on vérifie immédiatement qu'elle vérifie $I_s \geq -1$).

   Si deux espèces ne sont pas corrélées, $I_s = 0$. Dans le cas contraire, l'hypothèse de réacteur homogène peut conduire à surestimer ou sous-estimer la production chimique effective *à l'échelle du modèle*. Le cas classique est, pour l'application atmosphérique, donné par la réaction clé pour la pollution photochimique

$$NO + O_3 \rightarrow NO_2 + O_2 \tag{1.15}$$

   La situation classique est que le monoxyde d'azote NO est émis en limite inférieure de la couche limite (environ 90% des émissions d'oxydes d'azote sous cette forme),

tandis que l'ozone O_3, espèce à durée de vie plus longue, transportée et formée sur de longues distances domine au sommet de la couche limite. On va donc typiquement se trouver dans la situation où $I_s < 0$ et l'hypothèse homogène va revenir à surestimer la production effective par cette réaction chimique (en réalité, NO et O_3 ne sont pas vraiment en contact de manière homogène ...).

L'approximation n'est en particulier plus vérifiée pour les réactions les plus rapides au voisinage des sources fixes.

Sur la base de ces hypothèses simplificatrices, l'équation de dispersion moyennée devient donc :

$$\frac{\partial \langle c_i \rangle}{\partial t} + \mathrm{div}(\langle V(x,t) \rangle \langle c_i \rangle) = \mathrm{div}(\langle \rho \rangle K_{turb} \nabla \frac{\langle c_i \rangle}{\langle \rho \rangle}) + \chi_i(\langle c \rangle, \langle T(x,t) \rangle, t) + \langle S_i(x,t) \rangle \quad (1.16)$$

Dans toute la suite, on omettra de noter $\langle \Psi \rangle$ pour alléger les notations et on écrira Ψ.

1.1.4 Conditions aux limites

A cette équation d'Advection-Diffusion-Réaction sont associées des conditions initiales et des conditions aux limites.

Illustrons des conditions aux limites classiques rencontrées dans le cas atmosphérique. Une hypothèse usuelle revient à considérer que les phénomènes d'advection par le vent sont prépondérants horizontalement alors que les phénomènes de transport vertical sont dominés par la diffusion turbulente (brassage convectif de type Rayleigh-Bénard). Les conditions aux limites latérales sont donc les conditions aux limites classiques pour des problèmes hyperboliques (vent entrant), alors que les conditions au sol et au sommet du domaine considéré sont les suivantes, z désignant la coordonnée verticale :

1. **Au sol** ($z = 0$) :

$$-K_{turb}(x,t)\frac{\partial c_i}{\partial z} = E_i(x,t) - v^i_{dep} c_i \quad (1.17)$$

$E_i(x,t)$ est le terme d'émission surfacique de l'espèce i : il dépend du type de scénario d'émission choisi (rural, urbain, régional) et comprend une part d'origine anthropique liée au trafic routier et une part d'origine naturelle.

v^i_{dep} correspond à la vitesse de dépôt sec et est paramétrisée, par espèce chimique, en fonctions des conditions météos en couche limite et du type de sol (LUC : Land Use Coverage), à chaque type de sol correspondant une rugosité.

Mathématiquement parlant, ceci correspond à une condition de Robin.

2. **En sortie de couche limite** ($z = z_H$) :

$$-K_{turb}(x,t)\frac{\partial c_i}{\partial z} = 0 \quad (1.18)$$

qui correspond à la condition usuelle d'*atmosphère libre*. Mathématiquement parlant, ceci correspond à une condition de Neumann.

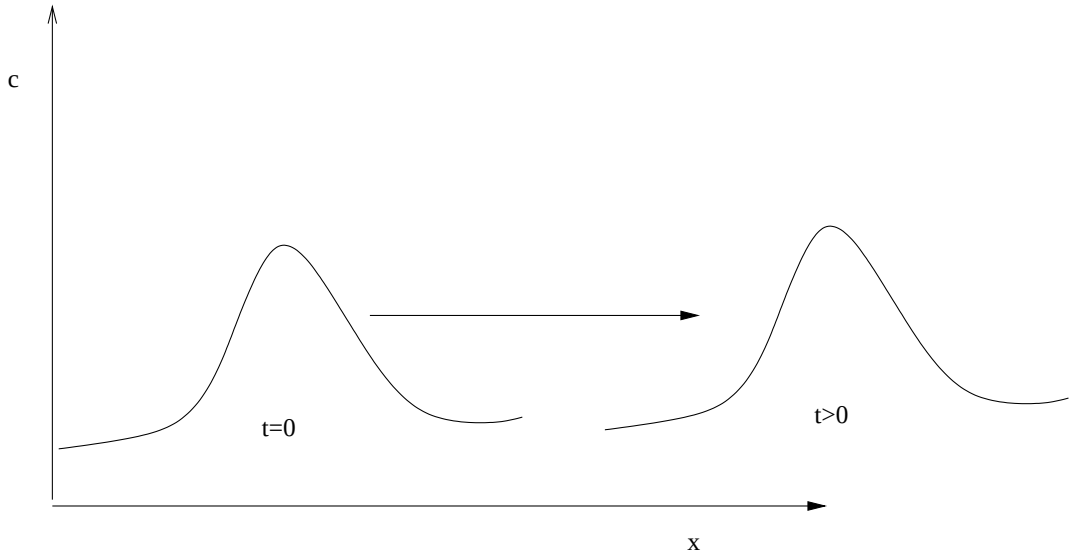

FIG. 1.2 – Advection.

1.2 Classification des processus

L'objectif de cette section est de rappeler brièvement la classification mathématique des processus qui ont été décrits dans l'équation de dispersion. L'intérêt d'une telle classification est par la suite de pouvoir recourir aux schémas numériques adéquats. En pratique (voir chapitre 2), on utilise des méthodes de séparation d'opérateurs qui reviennent à résoudre de manière découplée les processus décrits.

1.2.1 Advection

L'advection par le champ de vitesse V est donnée par :

$$\frac{\partial c_i}{\partial t} + \mathrm{div}(V(x,t)c_i) = 0 \tag{1.19}$$

Cette équation relève de la classe des *problèmes hyperboliques linéaires* (figure 1.2).

Un point clé associé à ces systèmes est bien sûr la vitesse de propagation de l'information (liée au champ de vitesse V). On se réfère au chapitre 4 pour les problématiques classiques associées (diffusion numérique, conditions de stabilité, ...).

1.2.2 Diffusion

L'équation de diffusion turbulente est donnée (pour une densité ρ constante) par :

$$\frac{\partial c_i}{\partial t} = \mathrm{div}(K_{turb}\nabla c_i) \tag{1.20}$$

Cette équation relève de la classe des *problèmes paraboliques* (figure 1.3). Les propriétés de cette équation (caractère "lissant") font que son intégration n'est pas en général un enjeu numérique en soi.

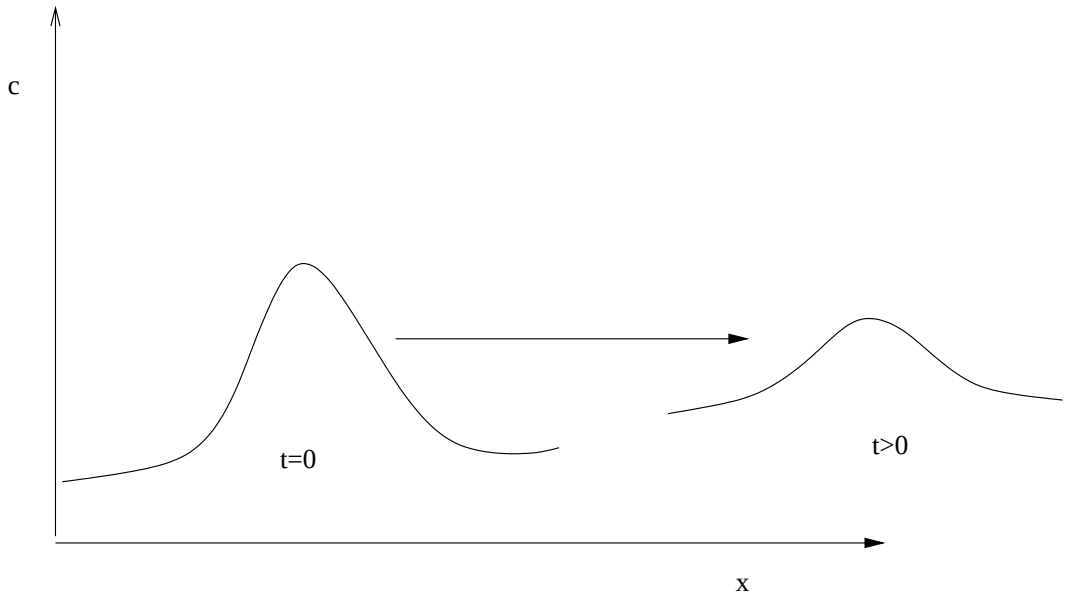

FIG. 1.3 – Diffusion.

1.2.3 Réaction

Les réactions chimiques sont décrites par :

$$\frac{dc_i}{dt} = \chi_i(c, T(x,t), t) \tag{1.21}$$

qui est un système d'Equations Différentielles Ordinaires (EDO). On se réfère au chapitre 3 pour les problématiques associées (systèmes "raides" -stiff-, schémas explicites/implicites, stabilité et positivité, réduction de modèles, etc).

Il est à noter que les processus d'advection et de diffusion ne couplent pas les espèces chimiques. Autrement dit, ces processus peuvent être résolus de manière parallèle sur toutes les espèces. A l'inverse, le terme réactif couple les espèces mais peut être résolu de manière parallèle sur toutes les mailles.

1.3 Discrétisation spatiale

Schématiquement, le modélisateur va avoir le choix entre les deux grandes classes usuelles de méthodes de discrétisation numérique :

1. les méthodes de type "éléments finis" qui reviennent à chercher les solutions $c(x,t)$ sous la forme $\sum_i \tilde{c}_i(t)u_i(x)$ avec (u_i) une base de fonctions prédéfinies dans un espace fonctionnel donné (typiquement des fonctions polynômiales à support spatial localisé). Les inconnues sont alors les composantes $\tilde{c}_i(t)$ qui peuvent être obtenues par une formulation faible de l'équation de départ. Modulo une troncature (via une projection dans un sous-espace de dimension finie), cette approche a le mérite de donner un cadre

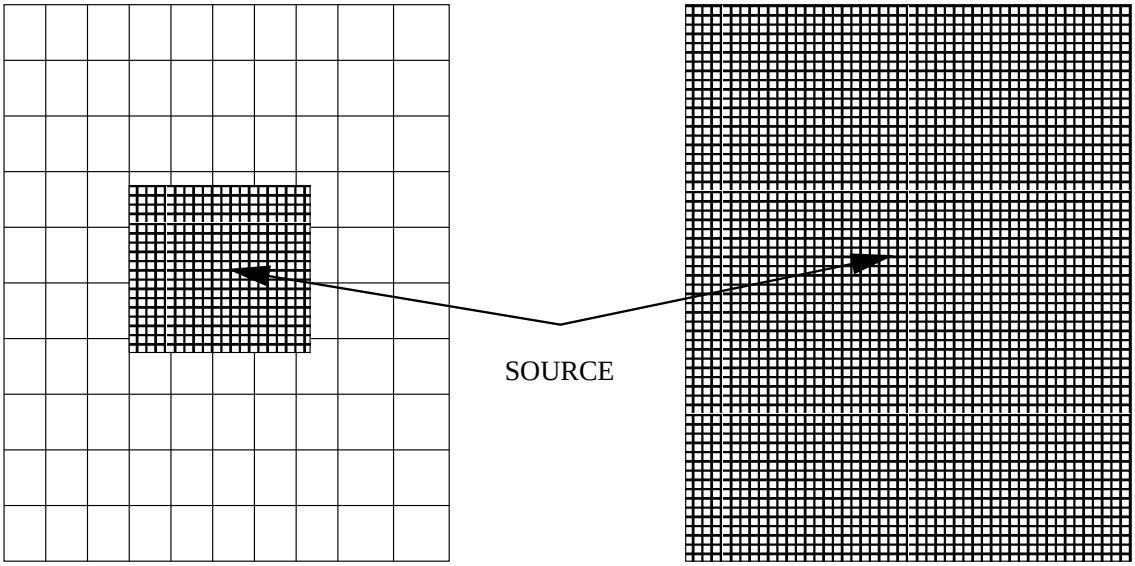

FIG. 1.4 – Nesting autour d'une source de pollution : imbrication de deux maillages (un grossier et un fin).

"fonctionnel" clair aux solutions obtenues. D'autre part, la prise en compte des singularités de l'écoulement (comme la présence d'une source par exemple autour de laquelle on souhaiterait avoir un raffinement de la solution) est aisée.

2. les méthodes de type "différences finies/volumes finis", plus simples à mettre en oeuvre (les variables discrétisées sont des valeurs de concentrations en des points de maillage ou des valeurs moyennes sur des mailles) mais dont l'inconvénient est la prise en compte des singularités. Une approche couramment utilisée est fournie par les méthodes d'imbrication de maillage (nesting) qui consistent à calculer la solution sur une hiérarchie de maillage associés à des domaines imbriqués (du plus fin autour de la source au plus grossier à grande échelle, figure 1.4), l'échange d'information entre les maillages se faisant selon plusieurs méthodes possibles.

De manière générale, la plupart des codes actuels du domaine ont recours à la méthode des différences finies/volumes finis, qui est celle que nous avons donc choisi de présenter.

1.4 Annexe : description du terme réactif

On détaille brièvement le terme de production chimique χ dans l'équation de dispersion (1.21).

1.4.1 Définitions générales

On se placera pour simplifier dans le cadre d'un modèle en phase gazeuse. Le terme de production chimique χ est alors donné pour une cinétique chimique générale [1] (pour n_e espèces et n_r réactions) par :

$$\chi(c,T,P) = S\omega(c,T,P)$$

où S est la matrice de stœchiométrie de dimension $n_e \times n_r$ et ω est le vecteur des n_r vitesses de réaction ; T et P désignent respectivement la température et la pression.

Une réaction *élémentaire* (c'est à dire qui a lieu effectivement entre espèces présentes simultanément) r est la donnée d'un jeu de coefficients stœchiométriques pour les réactants $(s_{ir}^-)_{i=1,n_e}$ et pour les produits $(s_{ir}^+)_{i=1,n_e}$. Elle est définie par le symbole :

$$\sum_{i=1}^{n_e} s_{ir}^- X_i \rightleftarrows \sum_{i=1}^{n_e} s_{ir}^+ X_i$$

où X_i est le symbole de l'espèce i. Notons que, du fait de la réversibilité des processus collisionnels, une réaction *élémentaire* est *toujours* réversible.

Les coefficients stœchiométriques globaux pour la réaction r sont donnés par :

$$s_{ir} = s_{ir}^+ - s_{ir}^-$$

La vitesse de réaction dans le sens direct (respectivement indirect) ω_r^+ (respectivement ω_r^-) est donnée par la loi d'action de masse selon :

$$\omega_r^+(c,T,P) = k_r^+(T,P) \prod_{i=1}^{i=n_e} c_i^{s_{ir}^+}$$

(respectivement :

$$\omega_r^-(c,T,P) = k_r^-(T,P) \prod_{i=1}^{i=n_e} c_i^{s_{ir}^-} \quad)$$

où $k_r^+(T,P)$ et $k_r^-(T,P)$ désignent les constantes cinétiques directe et indirecte de la réaction. La vitesse de réaction est alors donnée par :

$$\omega_r = \omega_r^+ - \omega_r^-$$

En règle générale, la constante directe est donnée par la loi (empirique) d'Arrhénius :

$$k_r^+(T,P) = A\,T^B\,exp(-\frac{E_a}{RT})$$

avec A la constante préexponentielle, B le facteur exponentiel et E_a l'énergie d'activation ; R est la constante des gaz parfaits.

La loi de Van't Hoff donne à partir de considérations d'équilibre thermodynamique la valeur de la constante inverse selon :

$$\frac{k_r^+(T,P)}{k_r^-(T,P)} = K_r^{eq}(T,P)$$

où $K_r^{eq}(T,P)$ est la constante d'équilibre de la réaction.

[1] On néglige pour le moment toute dépendance directe en temps via les phénomènes de photolyse.

1.4.2 Forme production-consommation

Il est aisé de vérifier, en séparant les réactions dans lesquelles X_i joue respectivement le rôle de réactant et de produit, que le terme de production pour la concentration c_i peut se mettre sous la forme dite communément de "production-consommation" :

$$\chi_i(c) = P_i(c) - L_i(c)c_i \tag{1.22}$$

où P_i et L_i sont respectivement les termes (positifs ou nuls) de production et de consommation. Sous forme vectorielle :

$$\chi(c) = P(c) - L(c)c \tag{1.23}$$

où P est le vecteur de production et L la matrice diagonale (positive ou nulle) de consommation. Les conditions thermodynamiques (T, P) sont ici implicitement fixées.

Cette forme est abondamment utilisée pour la définition de schémas numériques spécifiques à la chimie (chapitre 3).

1.4.3 Quelques remarques complémentaires

Avec le formalisme précédent, on définit usuellement le temps caractéristique de l'espèce i comme :

$$\tau_i(c) = \frac{1}{L_i(c)} \tag{1.24}$$

qui dépend d'une manière générale des concentrations c. On reviendra sur une définition plus précise dans le chapitre 3.

On peut se référer à [60] pour l'étude mathématique des équations de la cinétique chimique. Un point essentiel (et pour le moins attendu) est la positivité des concentrations chimiques. L'examen de (1.22) montre en effet que lorsque la concentration c_i s'annule, sa dérivée en temps devient positive :

$$P_i(c) \geq 0$$

Ceci permet de conclure formellement que c_i ne peut pas devenir négative. En réalité, l'argument est un peu plus "fin" et on se réfèrera à [60] (où l'on utilise l'analycité de $c(t)$).

1.4.4 Vers le couplage avec d'autres phases de la matière

En réalité, se limiter à la phase gazeuse est souvent trop restrictif. Par exemple, dans le cas de la chimie atmosphérique, les mécanismes en phase hétérogène (à la surface de cristaux de glace dans les nuages stratosphériques polaires) jouent un rôle clé pour expliquer les cycles de catalyse de destruction de l'ozone stratosphérique). Pour ce qui concerne la chimie troposphérique :

- de nombreux phénomènes ont lieu en phase aqueuse (dans les nuages),
- la phase condensée de la matière (solide ou liquide) peut interagir avec la phase gazeuse et a son intérêt propre (suivi des aérosols ou particules pour leur impact sur la santé ou la modification des propriétés radiatives et photolytiques de l'atmosphère).

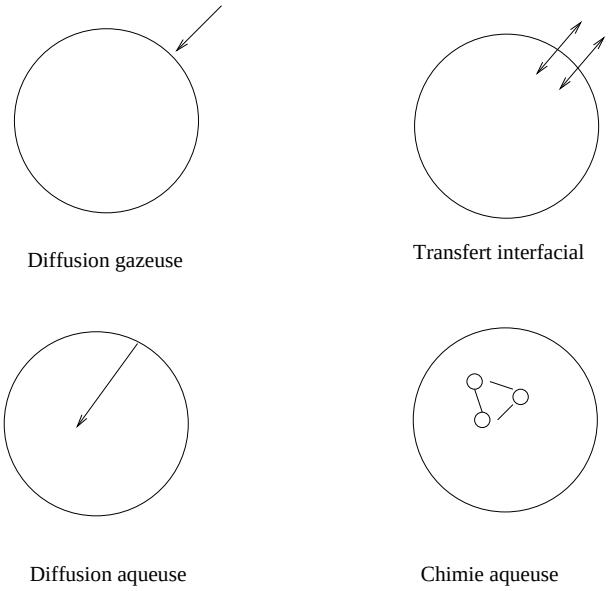

FIG. 1.5 – Transfert de masse au niveau d'une goutte de nuage

1.4.4.1 Modèle en phase aqueuse

Lorsque se produisent des épisodes nuageux, un transfert de masse a lieu entre la phase gazeuse des espèces et la phase aqueuse (les espèces dissoutes au sein des gouttes de nuages). Pour décrire ces processus, on a alors typiquement à prendre en compte (figure 1.5) :

- la diffusion moléculaire des espèces gazeuses vers les gouttes,
- le transfert interfacial à travers la surface de la goutte,
- la diffusion moléculaire des espèces dissoutes au sein des gouttes,
- les réactions chimiques en phase aqueuse.

Les trois premiers phénomènes relèvent du domaine de la *microphysique*. Pour mémoire, une goutte de nuage a une taille caractéristique de l'ordre de quelques dizaines de micromètres. Ces phénomènes sont importants pour le suivi de l'ozone régional, car ils peuvent constituer des puits d'ozone en phase gazeuse.

1.4.4.2 Aérosols et particules

On appelle aérosol la phase condensée de l'atmosphère, sous forme liquide ou solide. Les aérosols sont importants :

- pour eux-mêmes (impact sanitaire, surtout des nanoparticules) ;
- par la modification des propriétés radiatives de l'atmosphère (exemple : *effet direct* pour l'effet de serre) ;
- par l'interaction avec la phase gazeuse (au même titre que les gouttes de nuage) par les processus de condensation/évaporation ;
- par la modification des propriétés de formation des nuages par condensation de la vapeur d'eau sur des particules, les noyaux de condensation (CCN : Cloud Condensation Nuclei) ; on parle alors d'*effet indirect* pour l'effet de serre.

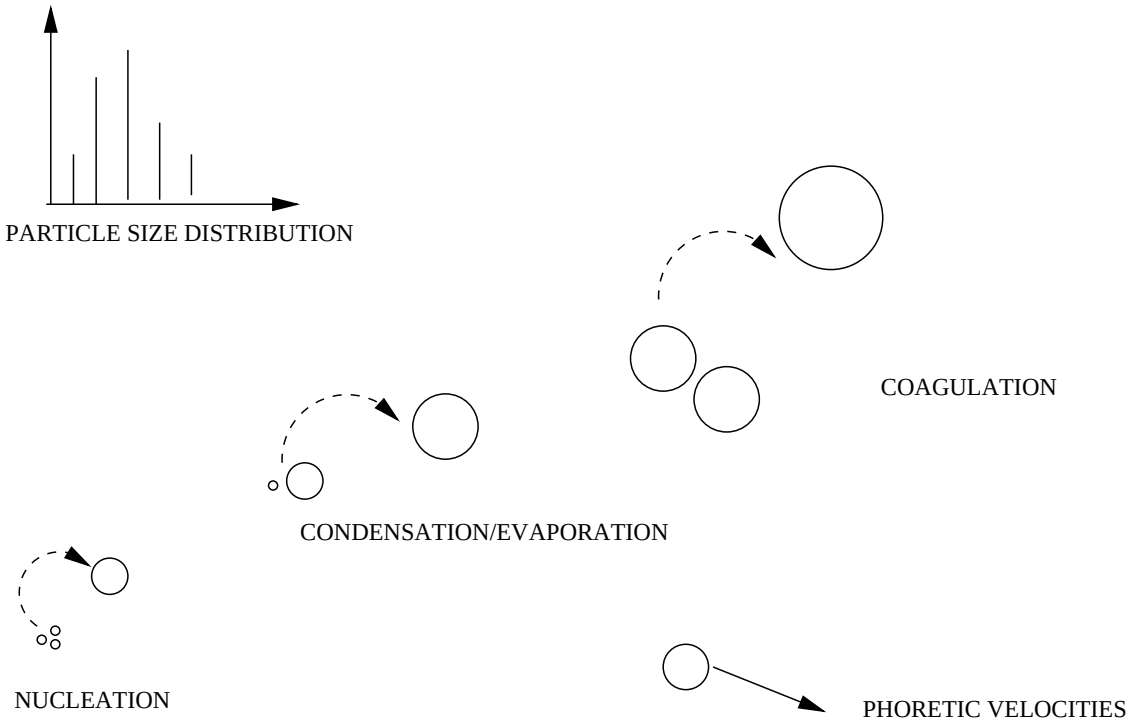

FIG. 1.6 – Les principaux processus affectant la dynamique des aérosols.

Il est hors de question ici de décrire la physique des aérosols (figure 1.6). On se contentera pour fixer les idées de préciser quelques points clés :

- la distribution en fonction de la taille des aérosols (supposés sphériques) est essentielle, car elle conditionne le dépôt des aérosols ;
- l'évolution de cette distribution est donnée par des modèles complexes, pour lesquels la question du renseignement des données (conditions initiales et paramètres) est déterminante.
  Si on appelle $n(v,t)$ la distribution d'un aérosol fixé en fonction du volume v à l'instant t, son évolution est donnée par l'équation de la dynamique des aérosols (GDE : General Dynamics Equation) :

$$\frac{\partial n}{\partial t} = \underbrace{\frac{1}{2}\int_{v_0}^{v} K(v-q,q)n(v-q,t)n(q,t)dq - n(v,t)\int_{v_0}^{\infty} K(q,v)n(q,t)dq}_{\text{coagulation}} \tag{1.25}$$

$$- \underbrace{\frac{\partial}{\partial v}(I(v)n(v,t))}_{\text{condensation-évaporation}} + \underbrace{J_0(v)\delta_{v_0}(v)}_{\text{nucléation}} + \underbrace{S(v)-R(v)}_{\text{puits et sources}}$$

où $K(.,.)$ est le noyau de coagulation (symétrique) et $I(v)$ est le taux de croissance par condensation et évaporation, qui est "piloté" par la thermodynamique. La *coagulation* décrit les phénomènes d'agrégation entre aérosols alors que les processus de *condensation-évaporation*

décrivent le gain ou la perte d'un *monomère* (l'aérosol de taille la plus petite) sous l'influence des conditions thermodynamiques. La *nucléation* précise les flux de création du plus petit aérosol pris en compte dans cette description continue (il est de volume v_0).

L'obtention de ce modèle continu à partir de la description initialement discrète de la population d'aérosols permet de définir exactement le noyau de nucléation.

Remarquons que ce modèle combine donc à la fois des termes intégro-différentiels (la coagulation) et hyperboliques (condensation-évaporation). C'est l'ensemble de ces termes qui jouent alors le rôle de terme réactif χ, ce qui rend d'autant plus difficile la résolution numérique ...

Chapitre 2

Méthodes de séparation d'opérateurs

Comme on l'a vu, l'équation de dispersion réactive met en jeu plusieurs processus (advection, diffusion, termes de pertes, termes de gain, termes réactifs, etc). De manière usuelle, l'ensemble de ces processus n'est pas résolu de manière couplée et des méthodes de "découplage" sont en pratique mises en oeuvre : on parle alors de *méthode de séparation d'opérateurs* (*operator splitting method*) ou de *méthode des pas fractionnaires* (*fractional step method*).

L'objet de ce chapitre est de présenter ces méthodes et l'analyse de leur comportement, notamment en terme d'évaluation des erreurs induites par le découplage.

On peut se référer par exemple à [27, 62] pour une présentation classique de ces méthodes.

2.1 Motivations

2.1.1 Notations

Dans le cas d'une densité constante (pour simplifier), l'équation de dispersion (à laquelle il faut bien entendu ajouter les conditions aux limites) :

$$\frac{\partial c_i}{\partial t} + \mathrm{div}(V(x,t)c_i) = \mathrm{div}(K(x,t)\nabla c_i) + \chi_i(c, T(x,t), t) + S_i(x,t) \tag{2.1}$$

peut être vue de manière générale comme une équation d'évolution :

$$\frac{dc}{dt} = \sum_{i=1}^{i=n_p} f_i(c) \tag{2.2}$$

mettant en jeu plusieurs processus $f_i(c)$, en nombre n_p. Dans le cas du modèle d'advection-diffusion-réaction, on a bien entendu :

$$f_1(c) = -\mathrm{div}(V(x,t)c) \ , \quad f_2(c) = \mathrm{div}(K(x,t)\nabla c) \ , \quad f_3(c) = \chi(c, T(x,t), t) \tag{2.3}$$

Les processus $f_i(c)$ sont donc à voir comme des *opérateurs* agissant sur les fonctions $c(.,t)$ de l'espace dans le cas du modèle continu. Après discrétisation spatiale éventuelle, ils seraient directement représentés par des fonctions agissant sur les vecteurs des valeurs ponctuelles de $c(.,t)$ (dans le cas des différences finies par exemple).

2.1.2 Méthode de séparation d'opérateurs versus résolution couplée

Sur le plan de la physique, l'ensemble des processus est bien entendu couplé et, en toute rigueur, les algorithmes numériques de résolution devraient résoudre de manière couplée les processus pris en compte.

Pour au moins deux raisons, une approche alternative de *séparation des opérateurs* est couramment mise en oeuvre :

1. en terme de modularité des codes informatiques résultants, on peut vouloir préférer utiliser une approche ne mettant en oeuvre que la résolution de processus pris indépendamment les uns des autres, ie :

$$\frac{dc}{dt} = f_i(c) \ , \quad c(0) = c_0 \tag{2.4}$$

Si on appelle $c^{(i)}(\Delta t, c_0)$ la solution du système précédent au temps $t = \Delta t$ (à voir également comme la sortie de l'appel d'une routine de résolution de ce système, c_0 étant la donnée d'entrée), les codes informatiques ne mettent alors en oeuvre que la résolution séquentielle de système du type (2.4). Un algorithme typique de résolution est alors pour une intégration sur un intervalle de temps $[0, T = N\Delta t]$:
DO n=1,N
c=c_n
DO i=1,n_p
c=$c^{(i)}(\Delta t, c)$
ENDDO
$c_{n+1} = c$
ENDDO

Un avantage clair est la grande modularité : l'ajout d'un nouveau processus n'affecte pas l'ensemble du programme et revient à ajouter une nouvelle "brique" résolvant un système de type (2.4) ; un processus peut être récupéré ou substitué auprès d'une autre équipe via l'incorporation de la brique concernée, etc.

2. en terme numérique, la résolution couplée peut également générer de nombreuses difficultés.

Les processus concernés peuvent avoir des comportements qualitatifs diamétralement opposés et les contraintes algorithmiques qui en résultent peuvent être difficiles à concilier : dans le cas découplé, on peut faire le choix de l'algorithme "optimal" pour chaque processus sans se soucier des autres processus.

De plus, dans le cas des modèles de dispersion réactive, le terme réactif se caractérise fréquemment par une grande complexité et une grande dimension (de nombreuses espèces concernées : dans le cas de la chimie atmosphérique, des centaines d'espèces chimiques traces sont ainsi modélisées). Une implication est alors souvent la grande dispersion des échelles de temps concernées (les temps caractéristiques couvrant plusieurs ordres de grandeurs). En anticipant sur la suite (chapitre 3 consacré à l'intégration en temps), ceci conduit à préférer des méthodes implicites d'intégration en temps. Dans le cas d'une méthode des lignes (Method of Lines, MOL : on discrétise d'abord en espace

puis en temps), une résolution couplée conduit alors à résoudre un système implicite de la forme :

$$\frac{c_{n+1}-c_n}{\Delta t} = \sum_{i=1}^{i=n_p} f_i(c_{n+1}) \stackrel{\Delta}{=} F(c_{n+1}) \tag{2.5}$$

où l'on notera que les processus ont été évalués au temps t_{n+1} (c_n est une approximation numérique de $c(t_n)$, $t_{n+1} = t_n + \Delta t$). L'équation algébrique en c_{n+1} doit alors être résolue et l'on verra que ceci passe par l'inversion de la matrice jacobienne de F. La taille de c est ici donnée par le produit du nombre de mailles (disons n_m) par le nombre d'espèces chimiques (disons n_e) : la complexité est alors de l'ordre de $O([n_e \times n_m]^3)$.

Dans le cas découplé, l'approche implicite ne sera utilisée que pour le processus présentant une grande dispersion d'échelles en temps (en l'occurence le terme réactif). On n'a alors plus à inverser une matrice que pour la résolution de ce processus (pour une variable de dimension n_e), dans les n_m mailles concernées. En supposant que cette étape est dimensionnante sur le plan calcul, la complexité est alors de $n_m \times O(n_e^3)$ qui est évidemment moindre que celle de l'approche couplée.

2.2 Analyse classique des méthodes de séparation d'opérateur dans le cas linéaire

La contrepartie des méthodes de séparation d'opérateurs est bien entendu l'erreur induite par le découplage des opérateurs. L'analyse se fait usuellement sur le cas linéaire que l'on présente dans un premier temps.

Dans toute la suite, on suppose obtenue la solution numérique c_n à l'itération n du splitting (ie après un temps $n\Delta t$, Δt étant ce que l'on appelle classiquement le *pas de temps de splitting*). On cherche alors à obtenir une approximation de c_{n+1} après intégration découplée des processus sur un intervalle de temps de longueur Δt. Les processus sont supposés être intégrés de manière exacte, éventuellement à l'aide de pas de temps inférieurs à Δt (on parle alors de *sous-cyclage*).

2.2.1 Méthode du premier ordre

On va considérer le problème d'évolution donné par deux processus linéaires représentés par A et B (des matrices dans le cas de la dimension finie après discrétisation) :

$$\frac{dc}{dt} = Ac + Bc \ , \quad c(0) = c_n \tag{2.6}$$

Soit Δt l'intervalle de temps de splitting. La méthode de séparation la plus naturelle est définie par les deux étapes successives :

1. Etape 1 de résolution du processus A :

$$\frac{dc^*}{dt} = Ac^* \quad \text{sur} \quad [0, \Delta t] \ , \quad c^*(0) = c_n \tag{2.7}$$

2. Etape 2 de résolution du processus B :

$$\frac{dc^{**}}{dt} = Bc^{**} \quad \text{sur} \quad [0, \Delta t] \ , \quad c^{**}(0) = c^{*}(\Delta t) \tag{2.8}$$

La valeur de $c(\Delta t)$ est alors approchée par $c^{**}(\Delta t)$. On appellera pour des raisons évidentes $(A - B)$ cet algorithme.

L'analyse classique de l'erreur induite par la séparation des opérateurs est effectuée à partir des solutions exponentielles en effectuant un développement asymptotique par rapport à l'intervalle de séparation Δt. Ici, la solution exacte est bien entendu :

$$c_{n+1} = \exp((A + B)\Delta t)c_n \tag{2.9}$$

alors que la solution calculée à l'aide de la méthode $(A - B)$ est

$$c_{A-B}(\Delta t) = \exp(B\Delta t)\exp(A\Delta t)c_n \tag{2.10}$$

L'erreur locale est alors :

$$le = c_{A-B}(\Delta t) - c_{n+1} = \frac{AB - BA}{2}\Delta t^2 c_n + O(\Delta t^3) \tag{2.11}$$

après développement des exponentielles. On a donc une méthode localement d'ordre 2 (globalement d'ordre 1) dans le cas général. Pour des opérateurs A et B qui commutent, il est clair que l'erreur de splitting est nulle.

2.2.2 Méthodes du second ordre

Il est bien sûr aisé de monter en ordre en notant que, pour le cas linéaire, le terme dominant de l'erreur est de signe opposé pour la méthode $(B - A)$ (obtenue par inversion de la séquence de résolution). Si l'on définit :

$$c_{n+1} = \frac{c_{A-B}(\Delta t) + c_{B-A}(\Delta t)}{2} \tag{2.12}$$

on a une solution d'ordre supérieur. Un inconvénient est le coût calcul, cet algorithme nécessitant la résolution de 4 processus élémentaires.

Strang ([48]) a proposé de symétriser la méthode précédente avec les trois étapes suivantes :

1. Etape 1 de résolution de B sur $[0, \Delta t/2]$:

$$\frac{dc^{*}}{dt} = Bc^{*} \quad \text{sur} \quad [0, \frac{\Delta t}{2}] \ , \quad c^{*}(0) = c_n \tag{2.13}$$

2. Etape 2 de résolution de A sur $[0, \Delta t]$:

$$\frac{dc^{**}}{dt} = Ac^{**} \quad \text{sur} \quad [0, \Delta t] \ , \quad c^{**}(0) = c^{*}(\frac{\Delta t}{2}) \tag{2.14}$$

3. Etape 3 de résolution de B sur $[\Delta t/2, \Delta t]$:

$$\frac{dc^{***}}{dt} = Bc^{***} \quad \text{sur} \quad [0, \frac{\Delta t}{2}] \ , \quad c^{***}(0) = c^{**}(\Delta t) \tag{2.15}$$

la valeur c_{n+1} étant approchée par $c^{***}(\frac{\Delta t}{2})$.

On nommera sans surprise $(B - A - B)$ cette méthode de séparation, dont un calcul immédiat permet de s'assurer qu'elle est bien d'ordre 3 localement (2 globalement) : la solution est de la forme

$$c_{B-A-B}(\Delta t) = \exp(B\Delta t/2) \exp(A\Delta t) \exp(B\Delta t/2) c_n \tag{2.16}$$

et un développement limité donne :

$$c_{B-A-B}(\Delta t) \simeq (I + B\Delta t/2 + B^2\Delta t^2/8)(I + A\Delta t + A^2\Delta t^2/2)(I + B\Delta t/2 + B^2\Delta t^2/8) c_n \tag{2.17}$$

qui tout calcul fait donne le développement limité à l'ordre 2 de la solution exacte.

Notons que le prix à payer n'est pas un intervalle d'intégration plus long (chaque opérateur est de toute manière intégré sur un intervalle de longueur Δt) mais deux interruptions d'intégration (et non plus une) [1].

2.2.3 Méthodes de type "Source Splitting"

Un inconvénient important des approches précédentes est le recours à une résolution séquentielle des processus : en pratique, les conditions initiales pour chaque processus sont modifiées à chaque sous-pas. Pour les processus présentant une grande disparité des échelles de temps, ceci conduit à éloigner les solutions intermédiaires des variétés d'équilibre associées aux temps caractéristiques lents (voir chapitre 3) et à intégrer de nombreuses phases transitoires générées de manière artificielle par l'approche séquentielle.

Une approche alternative est alors logiquement de ne pas modifier les conditions initiales mais de tenir compte des contributions des processus par des termes sources supplémentaires (on parle de *source splitting*) ou incréments (on parle aussi de *formulation incrémentale*) dans l'équation du second processus (celui qui présente la disparité d'échelles de temps et l'existence de phases transitoires potentielles).

L'algorithme devient alors le suivant :

1. Etape 1 de résolution du processus A (inchangée) :

$$\frac{dc^{*}}{dt} = Ac^{*} \quad \text{sur} \quad [0, \Delta t] \ , \quad c^{*}(0) = c_n \tag{2.18}$$

2. Etape 2 de résolution du processus B avec prise en compte d'un terme source (incrément) lié à l'étape 1 :

$$\frac{dc^{**}}{dt} = Bc^{**} + \frac{c^{*}(\Delta t) - c_n}{\Delta t} \quad \text{sur} \quad [0, \Delta t] \ , \quad c^{**}(0) = c_n \tag{2.19}$$

[1] A mettre en regard des remarques que l'on fera au chapitre 3 sur le coût des phases de redémarrage.

La valeur de c_{n+1} est alors approchée par $c^{**}(\Delta t)$.

Le point crucial est bien entendu que la condition initiale du second pas n'est pas modifiée. Autrement dit, si c_n avait atteint un état d'équilibre, on ne l'a pas perturbé (générant par là même une couche transitoire de relaxation vers cet équilibre).

Un développement limité montre que cette méthode est d'ordre 2 localement (1 globalement).

Pour être complet sur le plan de la terminologie, notamment pour les applications en océanographie et météorologie, on parle aussi de méthode d'intégration "par tendances" (la contribution du premier processus étant représentée par la "tendance" donnée par l'incrément).

2.2.4 Méthodes d'ordre supérieur

L'étude des erreurs d'ordre supérieur se fait à l'aide de la formule de Baker-Campbell-Hausdorff (BCH : voir [18, 24, 36]) pour deux opérateurs linéaires X et Y :

$$e^X e^Y = e^Z \ , \quad Z = X + Y + \frac{1}{2}[X,Y] + h(X,Y,[X,Y]) \tag{2.20}$$

avec $[X,Y] = XY - YX$ et $h(X,Y,0) = 0$.

Des méthodes d'ordre plus élevé peuvent également être obtenues à l'aide d'une simple extrapolation de Richardson ([55]). On rappelle que si l'on dispose d'un algorithme numérique définissant une solution $c_{\Delta t}$ pour un pas Δt, dont l'erreur locale est dominée par un terme du type $k\Delta t^2$, alors $(4c_{\Delta t/2} - c_{\Delta t})/3$ définit une solution d'ordre local 3. Le résultat est direct via :

$$c_{\Delta t/2} = c_{exact} + k\Delta t^2/4 \ , \quad c_{\Delta t} = c_{exact} + k\Delta t \tag{2.21}$$

avec c_{exact} la solution exacte.

Un inconvénient important de cette méthode simple à mettre en oeuvre est, d'une part son coût calcul, d'autre part la perte possible de positivité de la solution. En effet, même si les algorithmes mis en oeuvre pour la résolution permettent de garantir la positivité de $c_{\Delta t}$ pour tout pas de temps, ce n'est plus le cas pour la solution extrapolée.

2.2.5 Traitement des conditions aux limites

Dans les cas précédents, nous n'avons considéré que des opérateurs linéaires (ou ce qui revient au même des discrétisations spatiales de termes de transport *sans conditions aux limites*). La prise en compte de conditions aux limites induit le passage d'un cas linéaire à un cas affine.

Pour illustrer ce point, considérons par exemple le terme d'advection avec un schéma de type upwind (on se réfère au chapitre correspondant) : $c(x_i)$ est une approximation par différences finies de la valeur au point de maillage x_i (dans un cas 1D en espace) et après discrétisation du terme d'advection, on obtient comme équation différentielle le terme discrétisé générique :

$$\frac{dc(x_i)}{dt} = \frac{Vc(x_{i-1}) - Vc(x_i)}{\Delta x} \tag{2.22}$$

qui est bien sûr linéaire. Si l'on a une condition de bord de flux entrant, par exemple en $x = x_i$ selon $Vc(x_i) = L$ (avec L un terme de flux donné), alors le terme discrétisé devient affine.

La donne est la même pour le terme de diffusion, la condition aux limites de dépôt/émission générant un terme affine.

L'analyse précédente (limitée au cas linéaire) n'est alors plus valable et le traitement des conditions aux limites reste un problème largement ouvert pour les méthodes de séparation d'opérateurs (voir par exemple [42] ou [18] pour une quantification sur un exemple de réaction-advection).

Il est aisé de l'illustrer sur le cas affine :

$$\frac{dc}{dt} = Ac + Bc + L \tag{2.23}$$

où A et B sont des opérateurs linéaires qui commutent et L est un vecteur représentant les conditions aux limites. La question typique est de savoir comment "répartir" L entre les deux opérateurs, sachant qu'il n'y a pas d'erreur de splitting associée à la seule séparation des opérateurs. Pour le moment, on propose le schéma suivant :

$$\frac{dc^*}{dt} = Ac^* + \alpha L \quad \text{sur} \quad [0, \Delta t] \ , \quad c^*(0) = c_n \tag{2.24}$$

suivi de

$$\frac{dc^{**}}{dt} = Bc^{**} + \beta L \quad \text{sur} \quad [0, \Delta t] \ , \quad c^{**}(0) = c^*(\Delta t) \tag{2.25}$$

avec α et β deux coefficients à déterminer (plus exactement des matrices dans le cas vectoriel). On suppose que A, B et $A + B$ sont inversibles et on a directement pour la solution exacte :

$$c_{n+1} = e^{(A+B)\Delta t}(c_n + (A + B)^{-1}L) - (A + B)^{-1}L \tag{2.26}$$

et pour les solutions issues des deux étapes de splitting :

$$c^*(\Delta t) = e^{A\Delta t}(c_n + \alpha A^{-1}L) - \alpha A^{-1}L \tag{2.27}$$

et

$$c^{**}(\Delta t) = e^{B\Delta t}(c^*(\Delta t) + \beta B^{-1}L) - \beta B^{-1}L \tag{2.28}$$

soit encore :

$$c^{**}(\Delta t) = e^{(A+B)\Delta t}(c_n + \alpha A^{-1}L) + e^{B\Delta t}(\beta B^{-1}L - \alpha A^{-1}L) - \beta B^{-1}L \tag{2.29}$$

L'erreur de splitting est donc nulle si :

$$\alpha A^{-1} = (A + B)^{-1} \ , \quad \beta B^{-1} = (A + B)^{-1} \ , \quad \beta B^{-1} - \alpha A^{-1} = 0 \tag{2.30}$$

ce qui conduit au choix des paramètres α et β selon :

$$\alpha = (A + B)^{-1}A \ , \quad \beta = (A + B)^{-1}B \tag{2.31}$$

Notons que la répartition optimale des conditions aux limites n'est donc pas, *a priori*, donnée par l'origine physique des conditions.

Imaginons par exemple que L ne contienne que des conditions aux limites associées à l'opérateur A : on peut être tenté de prendre $\alpha = 1$ et $\beta = 0$. Néanmoins, un développement limité de l'erreur de splitting pour Δt petit permet de s'assurer que :

$$le = (1 - (\alpha + \beta))\Delta t + O(\Delta t^2) \tag{2.32}$$

et le schéma n'est donc que du second ordre dans ce cas-là.

Une telle approche se généralise aisément au cas non linéaire (voir par exemple [18]). Cependant, si de telles analyses permettent de trouver formellement un traitement des conditions aux limites, l'application en pratique d'une telle approche reste difficile.

2.2.6 Splitting au niveau de l'algèbre linéaire

Afin de s'affranchir des problèmes liés au traitement des conditions aux limites, une méthode alternative est d'effectuer le splitting "au niveau de l'algèbre linéaire" (certains auteurs parlent d' "internal splitting" : [54]). Dans le cadre d'une résolution couplée des opérateurs pour l'équation

$$\frac{dc}{dt} = Ac + Bc \ , \quad c(0) = c_n \tag{2.33}$$

l'utilisation d'un schéma implicite en temps conduit à la résolution d'équations algébriques (non linéaires dans le cas général). Par exemple avec une méthode d'Euler implicite :

$$\frac{c_{n+1} - c_n}{\Delta t} = Ac_{n+1} + Bc_{n+1} \tag{2.34}$$

soit $(I - (A + B)\Delta t)c_{n+1} = c_n$. L'idée est alors de faire par exemple l'approximation :

$$I - (A + B)\Delta t \sim (I - A\Delta t)(I - B\Delta t) + O(\Delta t^2) \tag{2.35}$$

ce qui conduit à la résolution successive de :

$$(I - A\Delta t)c^* = c_n \tag{2.36}$$

puis de

$$(I - B\Delta t)c^{**} = c^* \tag{2.37}$$

Le choix des notations n'est pas innocent et on reconnait la résolution par la méthode d'Euler implicite du splitting d'opérateurs $(A - B)$:

$$\frac{dc^*}{dt} = Ac^* \quad \text{sur} \quad [0, \Delta t] \ , \quad c^*(0) = c_n \tag{2.38}$$

suivi de

$$\frac{dc^{**}}{dt} = Bc^{**} \quad \text{sur} \quad [0, \Delta t] \ , \quad c^{**}(0) = c^*(\Delta t) \tag{2.39}$$

On peut se référer par exemple à [1, 54]. Notons qu'une méthode du même type est proposée pour la résolution des équations de Navier Stokes dans [30] (où les méthodes de splitting sont interprétées comme des choix de décomposition LU de matrices).

Notons pour conclure, que sur le plan de la terminologie, on parle aussi de méthode AMF (pour Approximate Matrix Factorization).

2.2.7 Extension au cas non linéaire

La généralisation au cas non linéaire de la notion de commutateur se fait avec l'utilisation de la *dérivée de Lie* associée aux fonctions f et g (de la variable c) :

$$[f, g] = \frac{\partial g}{\partial c} f - \frac{\partial f}{\partial c} g \tag{2.40}$$

Notons qu'à l'aide de la formule BCH, il suffit de montrer que deux opérateurs commutent pour être assuré que l'erreur de splitting est nulle.

2.3 Application au cas de l'équation d'advection-diffusion-réaction

2.3.1 Résultat

Les techniques précédentes ont été appliquées de manière systématique dans [24] aux équations d'Advection-Diffusion-Réaction (sans conditions aux limites) :

$$\frac{\partial c_i}{\partial t} + \mathrm{div}(V(x,t)c_i) = \mathrm{div}(K(x,t)\nabla c_i) + \chi_i(c, T(x,t), t) \tag{2.41}$$

L'ensemble des propriétés obtenues est résumé dans le théorème suivant.

Théorème 2.3.1

1. *L'advection et la chimie commutent si la chimie ne dépend pas de la position spatiale et si le champ de vitesse est à divergence nulle.*
2. *L'advection et la diffusion commutent si le champ de vitesse et la diffusion ne dépendent pas de la position spatiale.*
3. *La diffusion et la chimie commutent si la chimie est linéaire et ne dépend pas de la position spatiale.*

Donnons d'abord quelques commentaires. Il est clair que l'hypothèse de non dépendance en l'espace n'est pas en toute rigueur vérifiée pour la diffusion turbulente et le terme de chimie (via la température), mais on peut estimer qu'elle est valide localement. A l'inverse, une hypothèse "indéfendable" est celle relative à la linéarité de la chimie.

Ces résultats donnent donc une première indication : *la principale erreur de splitting est vraisemblablement celle liée au découplage entre diffusion (verticale) et chimie.*

Pour le cadre de la pollution atmosphérique, ceci conforte les observations de [13] et explique les efforts consacrés à une résolution couplée de la diffusion et de la chimie.

Une approche élégante de démonstration est fournie par le recours au formalisme de Lie mais la plupart des résultats peuvent être obtenus de manière certes un peu calculatoire mais peut-être plus parlante pour le lecteur non averti. C'est le choix qui a été retenu dans la suite pour la démonstration ([45]).

2.3.2 Advection-réaction

Il suffit d'écrire de manière classique la méthode des caractéristiques pour la commutation entre chimie et advection (comme noté dans [26] puis [19, 43]).

On suppose le champ de vitesse constant $V(x,t) = u$ et soit $X(x,t)$ la caractéristique associée à x : $X(x,t) = x + ut$. Soit la fonction de t à x fixé :

$$g_x(t) = c(X(x,t),t) \tag{2.42}$$

On a aisément :

$$\frac{dg_x}{dt} = \frac{\partial c}{\partial t} + \frac{\partial c}{\partial x}\frac{\partial X}{\partial t} = \chi(g_x,t) \ , \quad g_x(0) = c(x,0) \tag{2.43}$$

On notera désormais $G(g_0,t)$ la solution de l'EDO :

$$\frac{dG}{dt} = \chi(G) \ , \quad G(0) = g_0 \tag{2.44}$$

La solution du problème d'advection-réaction est donc donnée par :

$$c(x,t) = g_{x-ut}(t) = G(c(x-ut,0),t) \tag{2.45}$$

Montrons que cette solution est obtenue par splitting. Le splitting advection puis réaction revient à intégrer pour la seconde étape :

$$\frac{dc_x}{dt} = \chi(c_x,t) \ , \tag{2.46}$$

avec pour condition initiale $c_x(0) = c(x - u\Delta t, 0)$ la sortie en $t = \Delta t$ de l'étape d'advection. On a donc par définition, avec des notations évidentes :

$$c_{A-\chi}(x,t) = G(c(x-u\Delta t,0),t) \tag{2.47}$$

et on retrouve la solution exacte au temps $t = \Delta t$ [2].

Le splitting associé réaction-advection est donné directement par :

$$c_{\chi-A}(x,t) = G(c(x-ut,0),\Delta t) \tag{2.48}$$

où $G(c(x,0),\Delta t)$ est la condition initiale pour l'étape de convection. On retrouve bien entendu en $t = \Delta t$ la solution exacte.

2.3.3 Diffusion-Réaction

L'étude de la commutation des opérateurs de diffusion et de chimie linéaire peut être effectuée de la même manière à l'aide de la solution de l'équation de la chaleur.

On étudie le cas d'une diffusion indépendante de l'espace et scalaire. Pour simplifier, on prend la diffusion et la dimension de l'espace égales à 1. On rappelle enfin que la solution de

[2]... ce qui montre l'importance des temps de sortie des algorithmes de splitting : on ne retrouve la solution exacte qu'au temps $t = \Delta t$, les solutions intermédiaires n'ayant pas de "sens physique".

l'équation de la chaleur dans tout l'espace peut être calculée à l'aide du noyau $G(z,t)$ selon ([25]) :

$$c(x,t) = \int \quad G(x-y,t)\, c(y,0)\, dy \ , \quad G(z,t) = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{z^2}{4t}) \tag{2.49}$$

On note $\chi(c,t) = Mc$, où M est une matrice, le terme de chimie étant supposé linéaire. Notons enfin :

$$\bar{c}(x,t) = e^{-Mt} c(x,t) \tag{2.50}$$

où c est la solution de l'équation de Réaction-Diffusion initiale.

On a aisément :

$$\frac{\partial \bar{c}}{\partial t} = e^{-Mt} \frac{\partial c}{\partial t} - M e^{-Mt} c(x,t) \tag{2.51}$$

soit :

$$\frac{\partial \bar{c}}{\partial t} = e^{-Mt} \frac{\partial^2 c}{\partial x^2} = \frac{\partial^2 \bar{c}}{\partial x^2} \tag{2.52}$$

dont la solution est directement donnée par :

$$\bar{c}(x,t) = \int \quad G(x-y,t)\, c(y,0)\, dy \tag{2.53}$$

Finalement la solution exacte est :

$$c(x,t) = e^{Mt} \int \quad G(x-y,t)\, c(y,0)\, dy \tag{2.54}$$

dont l'évaluation en $t = \Delta t$ donne :

$$c(x,\Delta t) = e^{M\Delta t} \int \quad G(x-y,\Delta t)\, c(y,0)\, dy \tag{2.55}$$

Cette solution s'interprète directement comme la solution issue du splitting Diffusion puis Réaction. Il reste à étudier le splitting Réaction puis Diffusion. La solution de la seconde étape est :

$$c^{**}(x,t) = \int \quad G(x-y,t)\, c^{**}(y,0)\, dy \tag{2.56}$$

avec pour condition initiale $c^{**}(y,0) = e^{M\Delta t} c(y,0)$, et on conclut aisément.

2.3.4 Advection-Diffusion

Etudions à présent la commutation entre advection et diffusion. On se place dans le même cadre d'hypothèses simplificatrices pour la diffusion que précédemment. En définissant la solution le long des caractéristiques selon :

$$g(x,t) = c(x+ut,t) \tag{2.57}$$

on vérifie aisément que g vérifie l'équation de la chaleur, soit :

$$c(x,t) = \int \quad G(x-ut-y,\Delta t)\, c(y,0)\, dy \tag{2.58}$$

dont l'évaluation en $t = \Delta t$ donne :

$$c(x, \Delta t) = \int \quad G(x - u\Delta t - y, \Delta t)\, c(y, 0)\, dy \tag{2.59}$$

Ceci s'interprète directement comme la solution issue du splitting diffusion puis advection.

Si on effectue le changement de variables $Y = y + u\Delta t$, on obtient :

$$c(x, \Delta t) = \int \quad G(x - Y, \Delta t)\, c(Y - u\Delta t, 0)\, dY \tag{2.60}$$

et on retrouve la solution issue du splitting advection puis diffusion.

2.4 Exercices

2.4.1 Traitement des conditions aux limites

On considère le problème d'advection-réaction $\dfrac{\partial c}{\partial t} + \dfrac{\partial c}{\partial x} = c^2$, x variant dans $]0, 1[$, avec la condition de Dirichlet en $x = 0$:

$$c(0, t) = \frac{\sin^2(\pi t)}{1 - t \sin^2(\pi t)}$$

et avec la condition initiale :

$$c(x, 0) = \sin^2(\pi x)$$

La solution exacte est :

$$c(x, t) = \frac{\sin^2(\pi(x - t))}{1 - t \sin^2(\pi(x - t))}$$

On s'intéresse à l'intégration du premier pas de temps, i.e. sur $[0, \Delta t]$. *Intégrer la chimie puis l'advection. Comparer à la solution exacte.*

Calculer la solution issue de la séparation chimie puis advection. Comparer à la solution exacte.

2.4.2 "Source splitting"

Montrer que l'erreur locale de la méthode "source splitting" est d'ordre 2.

Chapitre 3

Simulation numérique des Equations Différentielles Ordinaires pour le traitement des termes réactifs

L'objet de ce chapitre est de présenter les bases de la simulation numérique des systèmes d'Equations Différentielles Ordinaires (EDO) qui sont associées aux termes réactifs dans le modèle de dispersion réactive.

Les méthodes de splitting d'opérateurs sont classiquement utilisées pour découpler la résolution des différents processus. Dans ce cadre, la résolution en temps, maille par maille, des termes réactifs est souvent extrêmement délicate et de loin la partie limitante dans les algorithmes numériques du fait de la grande disparité des échelles de temps.

Une première section est consacrée aux notions numériques classiques de ce domaine. En pratique, les problèmes rencontrés se caractérisent par la grande dispersion des échelles de temps : on parle classiquement de problèmes "raides" (*stiff*), qu'aborde la deuxième section. Dans la troisième partie, on présente quelques algorithmes classiques. Enfin, on termine par une section dédiée aux méthodes de réduction, qui consistent à modifier le système physique en "filtrant" les composantes rapides (c'est à dire celles qui génèrent la raideur numérique).

3.1 Quelques notions classiques

3.1.1 Système considéré

Dans toute la suite, on cherche à résoudre un système d'EDOs sous la forme :

$$\frac{dc}{dt} = f(c,t) \ , \quad c(0) = c_0 \ , \quad c \in \mathbf{R}^p \tag{3.1}$$

qui peut éventuellement provenir d'une discrétisation spatiale.

Pour des analyses de stabilité, on étudiera le système :

$$\frac{d\delta c}{dt} = \lambda \delta c \ , \quad c \in \mathbf{R} \tag{3.2}$$

avec δc une perturbation dans une direction (que l'on notera par abus c) et λ une valeur propre de la matrice jacobienne $J = \partial f/\partial c$.

Enfin, on suppose donnée une discrétisation du temps (t_n), avec un pas de temps constant pour simplifier Δt : $c_n \simeq c(t_n)$ est alors l'approximation numérique que l'on cherche à calculer de manière itérative. Les algorithmes vont donc consister à estimer c_{n+1} en fonction des estimations aux temps précédents.

3.1.2 Erreur locale, erreur globale et stabilité

Afin d'illustrer les notions d'erreurs et de stabilité, on va considérer l'exemple de la *θ-méthode* donnée par l'algorithme suivant :

$$\frac{c_{n+1} - c_n}{\Delta t} = (1-\theta) f(c_n, t_n) + \theta f(c_{n+1}, t_{n+1}) \tag{3.3}$$

où θ est un indicateur du degré d'implicitation : si $\theta = 0$ (resp. 1), on retrouve la méthode d'Euler *explicite* (resp. *implicite*).

Ecrivons à présent l'algorithme sous la forme :

$$c_{n+1} = c_n + \Delta t(1-\theta) f(c_n, t_n) + \Delta t \theta f(c_{n+1}, t_{n+1}) \tag{3.4}$$

Si on "insère" dans cette formule la solution exacte, on obtient :

$$c(t_{n+1}) = c(t_n) + \Delta t(1-\theta) f(c(t_n), t_n) + \Delta t \theta f(c(t_{n+1}), t_{n+1}) + \rho_n \tag{3.5}$$

avec un résidu ρ_n, la solution exacte n'ayant aucune raison de vérifier l'algorithme discret. Après développement de Taylor, on obtient :

$$\rho_n = \frac{1}{2}(1-2\theta)\Delta t^2 \frac{d^2 c}{dt^2} + O(\Delta t^3) \tag{3.6}$$

avec $\dfrac{d^2 c}{dt^2} = \dfrac{\partial f}{\partial c} f + \dfrac{\partial f}{\partial t}$. ρ_n définit ce que l'on appelle classiquement *l'erreur de troncature*.

Un point clé est bien entendu la propagation de cette erreur lors des pas de temps ultérieurs. Si on note $\varepsilon_n = c(t_n) - c_n$ *l'erreur globale* (résultant des erreurs locales antérieures et de leur propagation), on a pour le cas linéaire (linéarisé) :

$$\varepsilon_{n+1} = \varepsilon_n + (1-\theta)\lambda \Delta t \varepsilon_n + \theta \lambda \Delta t \varepsilon_{n+1} + \rho_n \tag{3.7}$$

En notant :

$$R(z) = \frac{1 + (1-\theta)z}{1 - \theta z} \tag{3.8}$$

l'étude de l'erreur devient :

$$\varepsilon_{n+1} = R(\lambda \Delta t)\varepsilon_n + \delta_n \tag{3.9}$$

avec $\delta_n = (1 - \theta\lambda\Delta t)^{-1}\rho_n$ *l'erreur locale* qui correspond à l'erreur effectuée sur un pas de temps, en supposant la solution bonne à t_n. Il y a donc de manière logique deux contributions à l'erreur :

1. une erreur purement locale ;
2. une erreur correspondant à la propagation des erreurs précédentes.

Avec une terminologie évidente, on appelle classiquement *fonction de stabilité* la fonction $R(z)$ de la variable complexe $z \in \mathbf{C}$ (les valeurs propres de J étant en toute généralité complexes) définie pour le problème linéarisé $f(c,t) = \lambda c$ par :

$$c_{n+1} = R(\lambda \Delta t) c_n \tag{3.10}$$

On a alors directement :

$$\varepsilon_n = (R(\lambda \Delta t))^n \varepsilon_0 + \sum_{i=0}^{i=n-1} (R(\lambda \Delta t))^i \delta_{n-1-i} \tag{3.11}$$

Si $||R(\lambda \Delta t)|| \leq K$ (fonction de stabilité bornée), alors on a :

$$|\varepsilon_n| \leq K|\varepsilon_0| + K \sum_{i=0}^{i=n-1} |\delta_i| \tag{3.12}$$

Pour une erreur locale $\delta_i = O(\Delta t^{p+1})$ ($p = 1$ ou $p = 2$ ici), on a donc avec $T = n\Delta t$, une erreur contrôlée, hors la partie relative aux conditions initiales, en $O(\Delta t^p)$, ce qui illustre la perte d'ordre lors du passage de l'erreur locale à l'erreur globale.

3.1.3 Domaines de stabilité

Pour le problème linéarisé, lorsque $\lambda \leq 0$, la perturbation décroît en valeur absolue. On peut attendre la même propriété (ce qui va au delà d'attendre que la perturbation soit bornée) pour le cas discrétisé. On définit donc assez logiquement les méthodes $A-$*stables* pour lesquelles $\mathbf{C}^-$ est contenue dans le domaine de stabilité :

$$S = \{z \in \mathbf{C} : |R(z)| \leq 1\} \tag{3.13}$$

On peut montrer que la θ-méthode est A-stable pour $\theta \geq 1/2$. Sinon, il est suffisant de vérifier que $\lambda \Delta t \in S$ pour garantir la stabilité de l'algorithme, ce qui donne une contrainte sur le pas de temps, en pratique pour le cas de la méthode d'Euler explicite :

$$\Delta t \leq \frac{1}{2|\lambda|} \tag{3.14}$$

Comme le concept de A-stabilité peut s'avérer trop restrictif, un concept moins fort et habituellement demandé aux schémas numériques est la $A(\alpha)$-*stabilité*, définie par $\{z : |arg(-z)| \leq \alpha\} \subset \mathbf{S}$.

Enfin, pour terminer avec ces concepts, il est fréquemment demandé (notamment pour les systèmes raides, voir ci-après) la *L-stabilité* qui est associée à $R(\infty) = 0$. En pratique, ceci correspond au cas asymptotique $\lambda \Delta t \gg 1$, dans lequel on souhaite travailler avec nos schémas numériques (pas de temps non contraints par les échelles physiques les plus petites).

valeur propre	espèce	τ^{-1}
-80019.18	O^3P	-80019.17
-78.34	$RXPAR$	-78.34
-54.68	OH	-51.41
-19.74	PHO	-19.74
-18.95	C_2O_3	-18.08
-17.98	NO_3	-17.98
-6.69	HO_2	-6.69
-5.44	XO_2	-9.18
-0.56	XO_2N	-0.56
-0.47	HNO_4	-0.85
-4.74E-2	N_2O_5	-5.25E-2
-3.10E-2	NO	-1.04E-2 (-3.16E-2)
	NO_2	-5.90E-3 (-3.14E-2)
	O_3	-1.51E-2 (-1.9E-2)

TAB. 3.1 – Raideur du schéma CBMIV.

Il est enfin à noter que hors "cas pathologique" (comme les systèmes autocatalytiques), la cinétique chimique est stable et les valeurs propres de la matrice jacobienne sont de partie réelle négative.

Un point clé est bien entendu le respect de la positivité des concentrations. En effet, pour le cas linéaire :

$$\frac{dc}{dt} = \lambda c \ , \quad \lambda \leq 0 \tag{3.15}$$

une concentration négative peut conduire à une "explosion" non contrôlée de la simulation numérique. Le critère de positivité va donc être un élément essentiel des schémas numériques.

3.2 Systèmes raides

3.2.1 Quelques caractérisations de la raideur

3.2.1.1 Distribution des valeurs propres et des temps caractéristiques

Une caractéristique essentielle des systèmes réactifs à traiter en pratique est la grande disparité des échelles de temps. Par exemple, pour le cas de l'atmosphère, les échelles varient de quelques millisecondes pour des radicaux comme OH ou O à des années pour le méthane.

Le tableau 3.1 montre, pour un schéma cinétique couramment utilisé en pollution atmosphérique ($CBMIV$ [11]), la distribution des 12 plus grandes valeurs propres (en valeur absolue) du Jacobien associé à la production chimique (en un point donné de l'espace des phases). Toutes les autres valeurs propres, au nombre de 16, sont supérieures ou égales à $-8.4.10^{-4}$, et se distribuent de manière continue jusqu'à la plus grande valeur propre ($-5.9.10^{-7}$).

La colonne en regard indique les temps caractéristiques (remarque 1.4.3) de certaines espèces. Les espèces NO_2 et O_3 ont été rajoutées. Pour ces deux dernières espèces et NO, un temps caractéristique a été également calculé dans la base de variables lumpées $O_x = O_3 + NO_2$ et $NO_x = NO + NO_2$ (qui remplacent respectivement O_3 et NO_2).

Au sens généralement admis du terme, les équations de la cinétique chimique sont donc *raides* puisque :

- les valeurs propres du jacobien sont de partie réelle strictement négative ;
- si λ_{min} et λ_{max} sont les valeurs propres respectivement de plus petite et de plus grande valeurs absolues [1], alors :

$$|\frac{\lambda_{max}}{\lambda_{min}}| \gg 1 \tag{3.16}$$

La notion de raideur est en réalité particulièrement difficile à définir [2], et on va d'abord revenir sur quelques caractéristiques des problèmes raides à l'aide de trois éclairages sensiblement différents :

- la perte de stabilité des schémas explicites,
- la comparaison des contraintes de précision et de stabilité,
- et enfin la résolution des systèmes algébriques induits par l'utilisation de schémas implicites.

On reviendra sur les différences entre schémas explicites et implicites en adoptant le point de vue des systèmes dynamiques dans la section consacrée à la réduction.

3.2.1.2 Stabilité versus précision

Une manière pragmatique de distinguer un problème raide d'un problème non raide se fonde sur la comparaison des performances d'un schéma explicite et d'un schéma implicite de même ordre ([4, 14, 39]). Ceci revient à dire qu'un problème est raide lorsque le pas de temps d'un schéma explicite est donné par la contrainte de stabilité plutôt que par la contrainte de précision. Avec une telle définition, la position dans l'intervalle de calcul joue un rôle prépondérant, comme l'exemple suivant va l'illustrer.

Examinons l'EDO (scalaire) [3] :

$$\frac{dc}{dt} = -\lambda(c - c_{eq}(t)) + \frac{dc_{eq}}{dt} \tag{3.17}$$

où $c_{eq}(t)$ est une fonction régulière connue (décrivant en réalité les modes "lents" du système) et $\lambda > 0$ un paramètre donné (essentiellement grand). La solution est bien entendu

$$c(t) = c_{eq}(t) + (c(0) - c_{eq}(0))e^{-\lambda t} \tag{3.18}$$

Il n'y a qu'une valeur propre et la notion de raideur doit être précisée.

---

[1] Ceci correspond en réalité au cas *bien partitionné* ("stiffly separable" dans [61]), que l'on étudiera en pratique.

[2] Voir la préface de [14] ou la discussion pages 360-363 dans [2].

[3] C'est une légère modification de l'exemple historique de Curtis et Hirschfelder ([8]) et bien entendu une variation sur l'exemple précédent (mais le point de vue est différent car $c_{eq}(t)$ est à voir comme un équivalent de solution "réduite" - voir plus loin).

Pour un schéma du premier ordre (comme les schémas d'Euler, implicite ou explicite), l'erreur de précision est classiquement contrôlée par un estimateur de la dérivée seconde de la solution :

$$|\rho_n| \simeq \frac{1}{2}|(1-2\theta)\Delta t^2 \frac{d^2c}{dt^2}| \leq \varepsilon_{tol} \tag{3.19}$$

avec ε_{tol} une tolérance d'erreur à spécifier par le modélisateur. Le point clé est que cette contraine est *la même* pour les deux schémas.

Montrons à présent que la contrainte de précision associée à cette erreur locale varie fortement en temps pour (3.17).

On a évidemment pour $t = O(\frac{1}{\lambda})$ (en temps court) :

$$c(t) \simeq c_{eq}(0) + (c(0) - c_{eq}(0))e^{-\lambda t} \tag{3.20}$$

La dérivée seconde de c est alors de l'ordre de :

$$\lambda^2(c(0) - c_{eq}(0))e^{-\lambda t} \tag{3.21}$$

et pour $\lambda \gg 1$, on obtient une contrainte de précision $(\lambda\Delta t)^2 \leq \varepsilon_{tol}$ beaucoup plus stricte que la contrainte de stabilité déjà calculée ($|\lambda\Delta t| \leq 1$).

A l'inverse, pour des temps assez grands, on a $c(t) \simeq c_{eq}(t)$, et la contrainte de précision est relâchée. Autrement dit, c'est la contrainte de stabilité qui va primer pour le schéma explicite.

On voit donc que l'on a deux intervalles en temps bien distincts :

- une phase transitoire (d'une durée de l'ordre de λ^{-1}), pour laquelle la contrainte de précision est du même ordre que la contrainte de stabilité du fait des gradients très prononcés. Un schéma explicite et un schéma implicite utilisent donc dans cette zone des pas de temps du même ordre.
- puis une phase dans laquelle la contrainte de précision devient non dimensionnante par rapport à la contrainte de stabilité d'un schéma explicite. C'est là qu'un schéma explicite a des performances dégradées. Ceci correspond donc à la partie *raide* de l'évolution [4].

Un corollaire pratique de cette remarque est que *les couches transitoires sont "chères" à intégrer* puisque la contrainte de précision y est stricte. Ceci permet de comprendre le coût des redémarrages et de l'intégration des phases transitoires pour les systèmes raides. Ceci est un point essentiel pour les méthodes de séparation d'opérateurs qui contribuent à créer des phases transitoires *artificielles* alors même que les phases transitoires *physiques* ont déjà été intégrées. C'est ce qui sous-tend le choix des méthodes de type "Source-Splitting" déjà présentées.

3.2.1.3 Dépendance aux conditions initiales

Pour l'exemple (3.17), afin d'étudier la dépendance en la condition initiale de $c(t)$, on note $C(t, c_0)$ la valeur, à l'instant t, de la solution issue de la condition initiale $c(0) = c_0$. On

[4] Le terme "raide" n'est donc pas à prendre au sens "alpin" (ou pyrénéen selon les affinités) du terme, puisqu'il suffit d'ajuster les conditions initiales avec $c(0) = c_{eq}(0)$ pour supprimer la phase transitoire (donc les gradients prononcés) et se placer directement dans la partie raide.

a alors directement :

$$\frac{\partial C(t, c_0)}{\partial c_0} = e^{-\lambda t} \tag{3.22}$$

et dès que la phase transitoire est passée ($t > \frac{1}{\lambda}$), on peut négliger la dépendance en fonction des conditions initiales :

$$\frac{\partial C(t, c_0)}{\partial c_0} \simeq 0 \tag{3.23}$$

Ceci se généralise pour un système avec plusieurs variables : les espèces *rapides* (celles concernées par les valeurs propres les plus négatives) ne dépendent pas des conditions initiales ! Les systèmes raides sont donc particulièrement stables.

De manière générale, pour sa composante rapide, le système "oublie" les conditions initiales et les erreurs accumulées (ce qui signifie en terme de données d'entrée des modèles par exemple, qu'il ne sert à rien d'estimer finement les espèces rapides !). On reviendra sur ce point lors de l'approche par réduction.

3.2.2 Mise en oeuvre pratique des algorithmes implicites

Pour les systèmes raides, hors phases transitoires initiales, il faut donc utiliser des méthodes implicites afin de ne pas être contraint par la stabilité.

L'utilisation de schémas implicites conduit néanmoins à des temps de calcul qui restent prohibitifs : en effet, même si les pas de temps sont plus grands que pour un schéma explicite, la résolution des systèmes algébriques associés aux schémas implicites reste coûteuse.

Pour illustrer ce point, prenons l'exemple (largement générique) de la méthode d'Euler implicite. Pour le système général initial, l'algorithme implicite s'écrit :

$$c_{n+1} = c_n + \Delta t f(c_{n+1}, t_{n+1}) \tag{3.24}$$

qui définit un système d'équations algébriques *a priori* non linéaires, qu'il s'agit de résoudre numériquement en l'inconnue c_{n+1}.

Une première approche simple est d'utiliser un algorithme de point fixe selon :

$$c_{n+1}^{(k+1)} = c_n + \Delta t f(c_{n+1}^{(k)}, t_{n+1}) \tag{3.25}$$

Ceci est simple à mettre en oeuvre puisque le calcul est ... explicite.

La convergence de la suite d'itérées $c_{n+1}^{(k)}$ doit donner c_{n+1} ... si la fonction $f(., t_{n+1})$ est contractante, c'est à dire que :

$$|\lambda| \Delta t \leq 1 \tag{3.26}$$

pour les λ valeurs propres de la matrice jacobienne de f. Autrement dit, on retrouve assez moralement la contrainte de stabilité des schémas explicites !

En pratique, on préfère chercher le zéro d'une fonction par une méthode itérative de type Newton (beaucoup moins restrictive pour la convergence : tout va dépendre de l'initialisation de la séquence, sur laquelle on ne s'étendra pas ici, même si c'est un enjeu numérique majeur). c_{n+1} est le zéro de la fonction :

$$g(c) = c - c_n - \Delta t f(c, t_{n+1}) \tag{3.27}$$

La méthode de Newton s'écrit alors :

$$(\frac{\partial g}{\partial c})_{c_{n+1}^{(k)}}(c_{n+1}^{(k+1)} - c_{n+1}^{(k)}) = -g(c_{n+1}^{(k)}) \tag{3.28}$$

Il s'agit donc d'inverser à chaque itération (et à chaque pas de temps ...) une matrice de dimension le nombre de traceurs (et ce en chaque maille). Heureusement, le fait de prendre une matrice approchée (en pratique fixée sur plusieurs itérations et sur plusieurs pas de temps) ne dégrade pas la solution (éventuellement la vitesse de convergence).

In fine, tout se ramène donc à une inversion de matrice, faite classiquement par une méthode de type décomposition LU.

3.3 Quelques algorithmes de résolution

On va à présent présenter plus spécifiquement quelques algorithmes de résolution de la cinétique chimique :

- les adaptations des méthodes multi-pas classiques, pour lesquelles se pose, comme on vient de le voir, la question de la résolution des systèmes algébriques induits,
- les méthodes *hybrides* de type "implicite-explicite",
- les méthodes *asymptotiques* qui se fondent sur la forme "production-consommation" (1.23) des équations de la cinétique chimique,
- et enfin les méthodes de Rosenbrock qui sont des méthodes particulièrement performantes (notamment pour la simulation de la pollution atmosphérique).

3.3.1 Méthodes multi-pas

L'algorithme de référence est la méthode de Gear (package LSODE [17]), qui est basée sur un schéma de type BDF (Backward Differentiation Formula). En dehors de l'avantage de pouvoir utiliser en "boîte noire" un tel logiciel, une méthode BDF permet en particulier de conserver les invariants linéaires ([31]). Un des désavantages majeurs est la non-positivité éventuelle des solutions et le recours, habituellement prôné, au *clipping* (la mise à zéro des concentrations négatives en cours de calcul).

La formule générale s'écrit sous la forme :

$$c_{n+1} = C_n + \beta \Delta t f(t_{n+1}, c_{n+1}) \tag{3.29}$$

où β est un coefficient et C_n est une combinaison linéaire des valeurs précédentes et des dérivées en ces points (tous les deux sont donc connus), qui dépendent de la méthode choisie.

Le potentiel de réduction du temps calcul d'une méthode BDF réside essentiellement dans l'accélération de la résolution de la contrainte algébrique :

1. soit par l'utilisation d'algorithmes d'algèbre linéaire dédiés, prenant en compte la faible densité des matrices traitées.

   Dans [20] la structure creuse des systèmes linéaires à résoudre permet l'utilisation conjointe d'algorithmes d'algèbre linéaire spécifiques et de techniques de vectorisation (sur l'ensemble des mailles), ce qui améliore considérablement les performances des

schémas BDF. Dans [33], on pourra trouver de même une analyse exhaustive des techniques de pivot (critère de Markowitz) pour minimiser la densité des systèmes linéaires à résoudre. Les résultats obtenus dans [53] (avec l'utilisation du logiciel VODE, dérivé de LSODE) vont dans la même direction.

2. soit par une résolution dégradée, alternative à l'algorithme de Newton.
   Une première approche se fonde sur la forme particulière des équations de la cinétique chimique, du moins en l'absence d'autocatalyse. Sous forme vectorielle, on a déjà vu qu'on avait une équation d'évolution de la forme (remarque 1.4.2) :

$$\frac{dc}{dt} = P(c) - L(c)c \tag{3.30}$$

   où c est le vecteur des concentrations, P est le vecteur (positif) de production et L est la matrice diagonale positive de consommation. Un schéma de type BDF appliqué à une telle équation différentielle conduit alors à la résolution du système algébrique réécrit sous la forme :

$$c_{n+1} = (I + \beta\Delta t L(c_{n+1}))^{-1}(C_n + \beta\Delta t P(c_{n+1})) \tag{3.31}$$

   où la matrice $I + \beta\Delta t L(c_{n+1})$ est diagonale et en conséquence inversible de manière directe.
   La méthode TWOSTEP, proposée dans [50, 51, 57], est fondée sur la méthode BDF d'ordre 2, avec une résolution par un algorithme de type Gauss-Seidel de l'équation algébrique :

$$c = (I + \beta\Delta t L(c))^{-1}(C_n + \beta\Delta t P(c)) \tag{3.32}$$

   En pratique, deux itérations suffisent pour résoudre cette équation. Une telle résolution, de type explicite (sans inversion de matrices), améliore considérablement les performances en terme de temps calcul (voir [57] et les benchmarks [35, 53]).

   Une seconde approche revient à approcher le Jacobien par une matrice triangulaire, ce qui évite là encore d'avoir à utiliser des méthodes itératives ou directes pour les inversions numériques ; elle est par exemple préconisée dans [23]. On rappelle qu'une approximation du Jacobien n'induit qu'une modification de la vitesse de convergence de l'algorithme de Newton ([7, 21]).

Remarque 3.3.1 (Un autre point de vue pour Twostep)
On peut voir Twostep selon un point de vue différent de celui de l'article initial ([51]) : pour schématiser, la méthode revient à préconditionner pour pouvoir utiliser un algorithme de point fixe.

L'algorithme de Newton appliqué à l'équation

$$g(c) \stackrel{def}{=} c - C_n - \beta\Delta t(P(c) - L(c)c) = 0 \tag{3.33}$$

s'écrit sous la forme :

$$\frac{\partial g}{\partial c}(c^k)(c^{k+1} - c^k) = -g(c^k) \tag{3.34}$$

avec

$$\frac{\partial g}{\partial c}(c^k) = I - \beta\Delta t \frac{\partial f}{\partial c}(c^k) \ , \quad f(c) = P(c) - L(c)c \tag{3.35}$$

Si on approche la matrice jacobienne par la matrice diagonale des termes de consommation selon :

$$\frac{\partial f}{\partial c}(c^k) \simeq -L(c^k) \tag{3.36}$$

on a aisément :

$$c^{k+1} = (I + \beta\Delta t L(c^k))^{-1}(C_n + \beta\Delta t P(c^k)) \tag{3.37}$$

et on retrouve exactement la formulation de Twostep.

On peut donc interpréter Twostep comme un choix d'approximation diagonale du Jacobien, lors de l'algorithme de Newton. Une telle approche a déjà été proposée par Shampine dans [38], avec notamment des conditions de convergence.

Ce point de vue permet de comprendre la nécessité d'utiliser des lumpings pour Twostep. Par exemple, pour le système (que l'on aura fréquemment l'occasion de rencontrer) :

$$\varepsilon\frac{dx}{dt} = -x + y \ , \quad \varepsilon\frac{dy}{dt} = x - y \tag{3.38}$$

le Jacobien est constant et bien entendu très mal approché par sa diagonale puisqu'en particulier on ne "voit" pas la valeur propre 0. Si on se place dans le système lumpé $(u = x+y, y)$, on a par contre :

$$\frac{du}{dt} = 0 \ , \quad \varepsilon\frac{dy}{dt} = u - 2y \tag{3.39}$$

et le Jacobien est bien approché par sa diagonale (au sens où les valeurs propres sont correctement restituées) [5]. ∎

Remarque 3.3.2 (Résolution couplée de la diffusion et de la chimie)
Les techniques utilisant le caractère creux des systèmes linéaires à résoudre ne sont pas généralisables aux cas 3D lorsque diffusion et chimie sont résolues de manière couplée (de préférence à une méthode de séparation d'opérateurs). L'algorithme Twostep présente par contre l'avantage de pouvoir être étendu au cas du couplage avec la diffusion ([52]) ∎

Remarque 3.3.3 (Méthode EBI)
La méthode EBI (Euler Backward Iterative) est fondée sur le schéma d'Euler implicite appliqué à l'équation :

$$\frac{dc}{dt} = P(c) - L(c)c \tag{3.40}$$

sous la forme :

$$c_i^{n+1} = f_i(c^{n+1}) = \frac{c_i^n + P_i(c^{n+1})\Delta t}{1 + L_i(c^{n+1})\Delta t} \tag{3.41}$$

qui est le pendant direct de Twostep (pour l'ordre 1). Une méthode de type point fixe

$$c_i^{n+1,k+1} = f_i(c^{n+1,k}) \tag{3.42}$$

[5] Ce qui est à mettre en regard du tableau 3.1.

ne converge pas, avec des pas de temps acceptables, pour des raisons qui ont déjà été évoquées.

La méthode EBI consiste alors à partitionner les espèces en plusieurs groupes G_1, G_2, ...G_{m-1} et G_m (selon la réactivité des espèces avec le radical OH et les termes de couplage existant entre espèces) et à résoudre par blocs de manière exacte le système algébrique, successivement pour chaque bloc G_i (avec $i < m$), et par point fixe pour le bloc restant G_m.

Des méthodes du même type sont proposées dans [10, 41]. On pourra également se référer à toute la littérature sur la "méthode des familles" ([5, 28]). ∎

3.3.2 Méthodes hybrides

A la suite des travaux de [3], des méthodes *hybrides* ont été proposées pour la résolution des équations de la cinétique chimique. L'idée directrice est de partitionner [6] les espèces en deux groupes, les espèces *lentes* x et les espèces *rapides* y, qui correspondent à une partition de la dynamique en une partie lente (f_L) et une partie rapide (f_R) selon :

$$\frac{dx}{dt} = f_L(x, y) \ , \quad \frac{dy}{dt} = f_R(x, y) \tag{3.43}$$

Un schéma explicite peut alors être appliqué pour la résolution de la partie non raide (pour l'intégration des espèces lentes) alors qu'un schéma implicite est utilisé pour la partie raide (correspondant aux espèces rapides).

L'avantage principal réside dans la diminution de la taille des systèmes à inverser, donnée à présent par le nombre de variables rapides. Les modes opératoires se distinguent pour les méthodes de ce type par le critère de partition, les schémas utilisés et l'ordre de succession des intégrations des sous-systèmes.

On peut par exemple se référer à [12] pour un algorithme fondé sur les schémas d'Euler implicite et explicite. La partition des variables se fait classiquement par comparaison des temps de vie des espèces et du pas de temps Δt. La séquence d'intégration des sous-systèmes proposée est la suivante :

– intégration du système non raide (à l'aide du schéma explicite) pour les espèces lentes,

$$\frac{x^*_{n+1} - x_n}{\Delta t} = f_L(x_n, y_n) \tag{3.44}$$

– intégration du système raide (à l'aide du schéma implicite) pour les espèces rapides, en utilisant les valeurs modifiées des espèces lentes,

$$\frac{y_{n+1} - y_n}{\Delta t} = f_R(x^*_{n+1}, y_{n+1}) \tag{3.45}$$

– réactualisation des espèces lentes (par intégration du système non raide tenant compte des espèces rapides modifiées) :

$$\frac{x_{n+1} - x_n}{\Delta t} = f_L(x_n, y_{n+1}) \tag{3.46}$$

[6] On ne précise pas plus ici.

Dans la même veine, la méthode IEH (Implicit Explicit Hybrid) utilise la méthode de Gear pour la partie raide (via le solveur LSODE) et une méthode de second ordre (Adams-Bashforth) pour la partie non raide ([6, 49]). La partition des variables n'est pas détaillée et au contraire de la méthode précédente la réactualisation des espèces lentes n'est pas effectuée.

On reviendra sur de telles approches dans la partie consacrée à la réduction .

3.3.3 Schémas asymptotiques

A la suite des travaux de Young et Boris (l'algorithme hybride CHEMEQ dans [63]), de nombreux schémas numériques à précision dégradée mais plus rapides [7] ont été proposés pour la simulation de la cinétique chimique.

Soit une espèce chimique i dont la concentration évolue selon :

$$\frac{dc_i}{dt} = f_i(c) = P_i(c) - L_i(c)c_i \tag{3.47}$$

avec P_i et L_i les termes (positifs ou nuls) de production et de consommation. Ces termes dépendent, en toute généralité, de l'ensemble des concentrations.

Les schémas asymptotiques sont fondés sur une hypothèse de *linéarisation* du terme source. En considérant en première approximation que P_i et L_i sont constants sur un intervalle de temps de longueur Δt, on a aisément :

$$c_i^{n+1} = \exp(-L_i^n \Delta t)c_i^n + (1 - \exp(-L_i^n \Delta t))\frac{P_i^n}{L_i^n} \tag{3.48}$$

en notant $c_i^n = c_i(t_n)$, $c^n = c(t_n)$, $L_i^n = L_i(c^n)$ et $P_i^n = P_i(c^n)$.

Le schéma QSSA (pour Quasi-Steady-State Approximation : [16]) revient à partitionner les espèces selon leur temps de vie $\tau_i^n = (L_i^n)^{-1}$ en un jeu d'espèces lentes et un jeu d'espèces rapides pour lesquels une intégration différente est réalisée. Notons que le nom est *trompeur*, car l'hypothèse essentielle est une hypothèse de linéarisation, sans lien avec une hypothèse de quasi-stationnarité (au sens de la section consacrée à la réduction).

La partition des espèces et la résolution de (3.48) sont faites par exemple de la manière suivante :

– pour les espèces lentes $\Delta t \leq \dfrac{\tau_i^n}{100}$, le schéma d'Euler explicite est utilisé :

$$c_i^{n+1} = c_i^n + (P_i^n - L_i^n c_i^n)\Delta t \tag{3.49}$$

– pour les espèces rapides $\Delta t \geq 10\tau_i^n$, (3.48) est approchée par :

$$c_i^{n+1} = \frac{P_i^n}{L_i^n} \tag{3.50}$$

[7] ... ou présumés comme tels, une analyse comparative du temps calcul à précision fixée se révèlant être particulièrement défavorable pour de tels schémas ([35, 37] par exemple).

– pour les espèces intermédiaires $\frac{\tau_i^n}{100} \leq \Delta t \leq 10\tau_i^n$, (3.48) est utilisée telle quelle.

Un tel algorithme est une version modifiée de l'algorithme de type prédicteur-correcteur initialement proposé par Young et Boris ([63]). Cette dernière méthode s'avère être moins rapide ([29]) mais plus précise ([37]). Notons que les implémentations varient considérablement d'un auteur à l'autre, ce qui rend particulièrement difficile toute tentative de comparaison ([37]).

Remarque 3.3.4 (QSSA d'ordre supérieur)
Plusieurs tentatives ont été faites pour améliorer la précision des méthodes de type QSSA. Une première approche, non spécifique aux méthodes QSSA, est fondée sur des extrapolations de Richardson ([9]) et permet d'améliorer considérablement la précision.

Une technique plus spécifique aux algorithmes de type QSSA est basée sur des développements d'ordre élevé de l'exponentielle ([57, 58]). Notons d'abord que la formule asymptotique (3.48) peut aussi s'écrire :

$$c_i^{n+1} = G(-L_i^n \Delta t) c_i^n + \Delta t \frac{G(-L_i^n \Delta t) - 1}{-L_i^n \Delta t} P_i^n \tag{3.51}$$

avec $G(z) = e^z$ ou bien une approximation consistante de l'exponentielle (par exemple un développement de Padé). La positivité des solutions est alors garantie pour

$$G(z) \geq 0 \ , \quad \frac{G(z) - 1}{z} \geq 0 \quad \text{pour } z \leq 0 \tag{3.52}$$

On pourra par exemple se référer à [58] pour la définition d'algorithmes QSSA d'ordre 2, fondés sur le développement de Padé :

$$G(z) = \frac{1}{1 - z + \frac{z^2}{2}} \qquad \blacksquare \tag{3.53}$$

Remarque 3.3.5 (A propos des règles "ad hoc")
Un inconvénient des schémas QSSA est la nécessité d'avoir recours à des règles *ad hoc* ("ad hoc rules" [8] dans [53]), souvent obscures.

On en relèvera notamment quatre :

– la résolution (simplifiée) pour les espèces rapides est faite chez certains auteurs par :

$$c_i^{n+1} = \frac{P_i^n}{L_i^n} \tag{3.54}$$

que l'on retrouve bien entendu si $\lim_{z \to -\infty} G(z) = 0$. Il est parfois préconisé ([15]) d'itérer un certain nombre de fois selon :

$$c_i^{n+1,k+1} = \frac{P_i^{n+1,k}}{L_i^{n+1,k}} \ , \quad k = 1, \ldots K \tag{3.55}$$

avec

$$P_i^{n+1,0} = P_i^n \ , \quad L_i^{n+1,0} = L_i^n \tag{3.56}$$

pour mieux prendre en compte les couplages entre espèces rapides.

[8] Le terme poli(tiquement correct) pour désigner l'huile de coude informatique.

– on notera que la contrainte algébrique (de fait) qui est utilisée pour calculer les espèces rapides en t_{n+1} est évaluée en t_n. Cela revient à décaler en temps la dépendance des espèces rapides vis à vis des lentes, alors que la bonne contrainte (voir section consacrée à la réduction) s'écrit :

$$c_i^{n+1} = \frac{P_i^{n+1}}{L_i^{n+1}} \tag{3.57}$$

C'est un des points essentiels qui explique la perte de précision d'une telle implémentation des méthodes QSSA.

– certains auteurs ([15]) préconisent d'ordonner les espèces rapides. Notons que, de la même manière, les espèces doivent être ordonnées pour la résolution du système linéaire par l'algorithme de Gauss-Seidel dans le cadre de Twostep ([57]).

– il est fréquemment recommandé ([15, 16, 40, 53, 57]) de travailler avec des espèces "lumpées", définies comme combinaison linéaire des espèces initiales ("the lumping trick" dans [50], page 81). Une telle technique permet d'améliorer notablement la précision des schémas QSSA. Les explications diffèrent grandement d'un auteur à l'autre : argument de conservation de masse pour certains groupes d'atomes ([15, 16, 29]) ou analogie avec un préconditionnement de système ([53]).
Les lumpings proposés varient également d'un schéma cinétique à un autre, ce qui est moral, mais d'une manière plus troublante également à schéma fixé.
Il est habituellement défini, sans plus de précision, les espèces suivantes ([16]) :

$$NO_x = NO + NO_2 \ , \quad O_x = NO_2 + O_3 \tag{3.58}$$

voire ([15])

$$O_3NO = O_3 - NO \ , \quad NO_z = NO_3 + N_2O_5 \tag{3.59}$$

$$NO_y = NO + NO_2 + NO_3 + 2N_2O_5 + HNO_2 + HNO_4 + PAN \tag{3.60}$$

et [9]

$$HO_x = OH + HO_2 \ , \quad PANX = PAN + C_2O_3 \tag{3.61}$$

En particulier, les techniques de lumping sont étroitement liées à la construction des modèles réduits (voir plus loin).

Un autre point essentiel concerne la résolution de la contrainte algébrique (3.57) : en quelques mots, les schémas QSSA sont peu précis, non pas du fait de l'hypothèse de quasi-stationnarité [10], mais parce que la contrainte algébrique est mal résolue numériquement. ■

3.3.4 Méthodes de type Rosenbrock

Les méthodes de type Rosenbrock fournissent des bons exemples de schémas rapides et à précision suffisante pour les applications de type pollution atmosphérique (une erreur relative en deça du %). L'idée générique ([32]) est de remplacer les systèmes non linéaires qui apparaissent dans les méthodes implicites directement par des systèmes linéaires qui

[9] On s'arrêtera là.

[10] Qui est de toute manière faussement sous-jacente.

ne sont plus construits *lors* de la résolution des systèmes non linéaires, par exemple avec l'algorithme de Newton, mais qui sont donnés, *dès le départ*, avec le schéma considéré [11]. C'est par exemple le cas de la méthode à un pas :

$$c_{n+1} = c_n + k \ , \quad (I - \Delta t \frac{\partial f}{\partial c}(c_n))k = \Delta t f(c_n) \tag{3.62}$$

pour la résolution de l'EDO

$$\frac{dc}{dt} = f(c) \tag{3.63}$$

Cette méthode correspond bien entendu à la première itération d'un algorithme de Newton utilisé pour la résolution du système algébrique issu d'un schéma d'Euler implicite.

Une telle méthode se généralise et on définit une méthode de Rosenbrock à s pas par :

$$c_{n+1} = c_n + \sum_{i=1}^{i=s} b_i k_i \ , \quad k_i = \Delta t f(c_n + \sum_{j=1}^{j=i-1} \alpha_{ij} k_j) + \Delta t \frac{\partial f}{\partial c}(c_n) \sum_{j=1}^{j=i} \gamma_{ij} k_j \tag{3.64}$$

les coefficients b_i, α_{ij} et γ_{ij} étant donnés pour chaque schéma (et fixés par des considérations de consistance essentiellement).

On se réfère à [14, 56] pour une présentation exhaustive des méthodes de Rosenbrock et leur application à la simulation de la pollution atmosphérique. Les méthodes, à respectivement deux et quatre pas, $ROS2$ et $RODAS3$ sont notamment testées avec succès dans [22, 44, 59]. Par contre, seule la méthode $ROS2$ conserve la propriété de positivité des solutions. Le benchmark [34] confirme que les méthodes de Rosenbrock sont actuellement les méthodes les plus efficaces en terme de compromis coût-précision pour la pollution atmosphérique.

Par exemple, la méthode ROS2 s'écrit :

$$c_{n+1} = c_n + \frac{1}{2}(k_1 + k_2) \tag{3.65}$$

$$(I - \gamma \Delta t J)k_1 = f(c_n, t_n) \ , \quad (I - \gamma \Delta t J)k_2 = f(c_n + \Delta t k_1, t_{n+1}) - 2\gamma \Delta t J k_1$$

avec J une approximation du Jacobien de f et $\gamma = 1 + 1/\sqrt{2}$ qui permet de garantir la L-stabilité de la méthode.

3.4 Réduction de modèles

Pour finir ce chapitre, on va présenter brièvement le cadre théorique qui sous-tend l'existence de la raideur numérique.

On a vu que les difficultés rencontrées pour la résolution numérique des systèmes raides proviennent de la grande disparité des échelles de temps. Modulo un adimensionnement et

[11] On parle aussi de méthodes *linéairement implicites*.

un changement éventuel de base, le système (que l'on a déjà introduit sous une forme analogue pour la présentation des méthodes hybrides) peut alors s'écrire sous la forme suivante (souvent appelée "lent/rapide") :

$$\frac{dx}{dt} = f_L(x,y) \ , \quad \varepsilon \frac{dy}{dt} = f_R(x,y) \tag{3.66}$$

avec x (resp. y) les espèces lentes (resp. rapides) et ε un rapport d'échelle de temps caractéristiques supposé être très petit par rapport à 1.

On peut alors montrer que pour de tels systèmes, il existe après une phase transitoire de durée $O(\varepsilon)$ un modèle *réduit* donné par :

$$\frac{dx}{dt} = f_L(x,y) \ , \quad f_R(x,y) = 0 \tag{3.67}$$

qui approche à $O(\varepsilon)$ près le système initial.

Ce système de dimension réduite (sa dimension est la dimension des variables lentes x) est un système algébro-différentiel défini par un système de contraintes algébriques $f_R(x,y) = 0$ définissant les variables rapides y en fonction des variables lentes x par une relation du type $y = h(x)$. Par application du théorème des fonctions implicites, ceci nécessite bien sûr que $\partial f_R / \partial y$ soit inversible (en réalité de valeurs propres à partie réelle strictement négative pour assurer la convergence en dehors de la couche transitoire vers ce modèle réduit).

Ce théorème (le théorème de Tikhonov ou le théorème de la variété centrale "globale") permet donc de donner un cadre à de nombreuses propriétés déjà rencontrées au cours de ce chapitre et dont un exemple très simple avait été fourni par l'exemple (3.17) : la non dépendance à des conditions initiales, la très grande stabilité du modèle hors couche transitoire (le modèle convergeant systématiquement vers la contrainte $y = h(x)$), l'intérêt de travailler dans des bases de variables spécifiques (les lumpings s'interprétant comme les changements de base permettant de partitionner le système sous la forme requise), etc.

Une alternative séduisante à l'utilisation de méthodes implicites est donc la construction puis la résolution des modèles réduits : les échelles de temps rapides ayant été filtrées, des algorithmes explicites peuvent alors être utilisés.

En cinétique chimique, les modèles réduits sont une généralisation des techniques classiques de type *Approximation de l'Etat Quasi Stationnaire* (AEQS ou QSSA en anglais) ou *Approximation de l'Equilibre Partiel*. Un exemple est par exemple fourni dans la figure 3.1 par la convergence pour plusieurs jeux de conditions initiales du système NO, NO_2 et O_3 vers le modèle réduit défini par l'équilibre de la réaction *globale* :

$$NO + O_3 \stackrel{1}{\rightarrow} NO_2 \ , \quad NO_2 \stackrel{2}{\rightarrow} NO + O_3 \tag{3.68}$$

que l'on ne détaille pas plus (le lecteur averti aura remarqué que ces réactions ne conservent même pas les éléments).

L'équilibre est alors donné par :

$$\frac{c_{NO} c_{O_3}}{c_{NO_2}} \simeq \frac{k_2}{k_1} \tag{3.69}$$

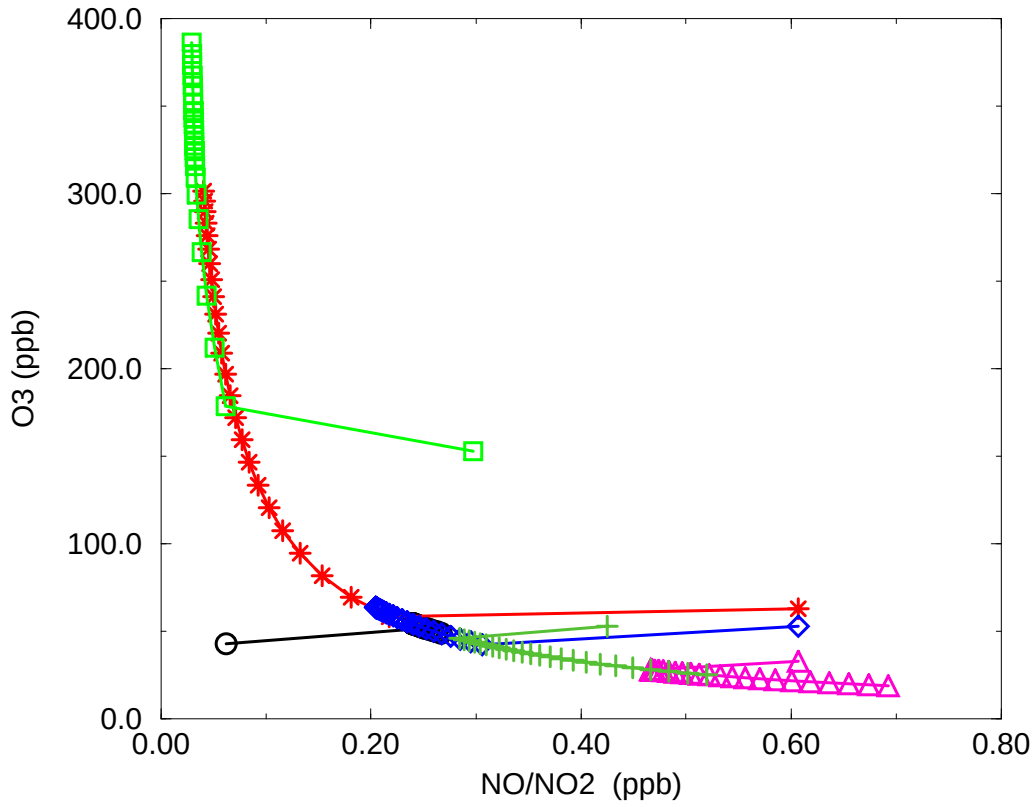

FIG. 3.1 – Convergence vers le modèle réduit pour le système NO, NO_2 et O_3.

avec k_i la constante cinétique de la réaction i.

Un écueil important à l'utilisation de ces méthodes est la difficulté de mise en oeuvre lorsque le modèle réduit ne peut pas être calculé explicitement ou doit être adapté à chaque pas de temps (par exemple, sa dimension variant au cours du temps). Par contre, ce cadre permet de donner un cadre extrêmement puissant pour comprendre le comportement qualitatif des modèles et des schémas numériques (par exemple pour les splittings).

Notons pour terminer ce chapitre que ces propriétés (systèmes multi-échelles générant une raideur numérique et associés à une hiérarchie de modèles) peuvent s'étendre à une large classe de système :

1. par exemple, pour la modélisation des aérosols, les dynamiques rapides correspondent à la mise à l'équilibre thermodynamique (entre phase gazeuse et phase particulaire) des petites particules submicroniques ;
2. en dynamique géophysique, les approximations quasi-géostrophiques pour le champ de vitesse (équilibré par les gradients de pression) rentrent aussi dans ce cadre théorique. Les problèmes étant de nature ondulatoire, les valeurs propres sont alors imaginaires pures le modèle générique étant :

$$\frac{dc}{dt} = j\lambda(c - c_{eq}(t)) + \frac{dc_{eq}}{dt} \ , \quad j^2 = -1 \ , \quad \lambda > 0 \tag{3.70}$$

Le modèle réduit correspond alors à un modèle "filtré" (ou moyenné) des composantes initiales des ondes à courte fréquence,

3.5 Exercices

3.5.1 Stabilité

On s'intéresse à l'équation :

$$\frac{\mathrm{d}c}{\mathrm{d}t} = \lambda c, \quad \lambda < 0 \tag{3.71}$$

Retrouver la condition de stabilité du schéma d'Euler explicite et montrer que le schéma d'Euler implicite est inconditionnellement stable.

Les méthodes d'Euler explicite et implicite sont des méthodes de Runge-Kutta, i.e. des méthodes qui, pour résoudre $c'(t) = F(t, c(t))$, se mettent sous la forme :

$$\begin{array}{c} c_{n+1} = c_n + \Delta t \sum_{i=1}^{s} \beta_i F(t_n + \gamma_i \Delta t, c_{n,i}) \\ \text{où} \qquad \forall i \in [\![1, s]\!] \quad c_{n,i} = c_n + \Delta t \sum_{j=1}^{s} \alpha_{ij} F(t_n + \gamma_j \Delta t, c_{n,j}) \end{array} \tag{3.72}$$

À quelle condition une méthode de Runge-Kutta sous la forme (3.72) est-elle explicite ?

Soit A la matrice $(\alpha_{i,j})_{i,j}$, b le vecteur $(\beta_i)_i$ et e le vecteur $(1, \cdots, 1)^T$ de longueur s. *Pour l'équation (3.71), écrire le schéma sous la forme $c_{n+1} = R(z)c_n$ où $z = \lambda \Delta t$.*

R est la fonction de stabilité. La région de stabilité est définie par $S = \{z \in \mathbb{C} / |R(z)| \leq 1\}$. *Déterminer la région de stabilité de la méthode d'Euler explicite.*

Une méthode est dite A-stable si $S \supset \mathbb{C}^- = \{z \in \mathbb{C} / \mathrm{Re}(z) \leq 0\}$. Elle est dite L-stable si $\lim_{z \to -\infty} |R(z)| = 0$. *Montrer que la méthode d'Euler implicite est L-stable.*

3.5.2 Retour sur le splitting

On souhaite montrer que, dans le cas raide, l'analyse classique de l'erreur due à la séparation d'opérateurs n'est plus valide. On considère l'équation raide :

$$c' = \frac{A}{\varepsilon}c + Bc \tag{3.73}$$

Calculer l'erreur due au splitting $A - B$.

Chapitre 4

Simulation numérique des processus d'advection-diffusion

Dans le cadre d'une méthode de séparation d'opérateurs, les processus d'advection et de diffusion sont résolus indépendamment des termes réactifs dans l'équation de dispersion réactive. Le point le plus difficile est en pratique l'advection (même lorsqu'elle est linéaire) car des propriétés qualitatives strictes doivent être respectées numériquement : conservation de la masse, positivité, monotonie.

Ce chapitre est organisé de la manière suivante. Dans la première section, on traite l'advection linéaire. Après avoir rappelé la base de l'approche lagrangienne (méthode des caractéristiques), on présente brièvement les notions classiques de stabilité. Le comportement qualitatif de la solution numérique est étudié à l'aide de la notion d'*EDP équivalente* qui permet notamment de préciser la *diffusion numérique* observée. Une partie spécifique traite des méthodes permettant de réduire la diffusion numérique (méthodes à limiteurs de flux), ce qui est en pratique l'enjeu principal de la simulation numérique de cette classe de problèmes.

Dans une seconde section, la simulation numérique de la diffusion est rapidement présentée, ne présentant pas d'écueil particulier.

4.1 Advection

4.1.1 Modèle. Méthode des caractéristiques.

On considère ici l'advection d'un traceur, de concentration c, dans un milieu de vitesse supposée connue et constante dans un premier temps, V :

$$\frac{\partial c}{\partial t} + \mathrm{div}(Vc) = 0 \ , \quad c(x, t = 0) = c_0(x) \tag{4.1}$$

avec des conditions aux limites si nécessaire (domaine non borné).

Dans le cas où V est constante, il y a bien entendu une solution évidente (figure 4.1) :

$$c(x, t) = c_0(x - Vt) \tag{4.2}$$

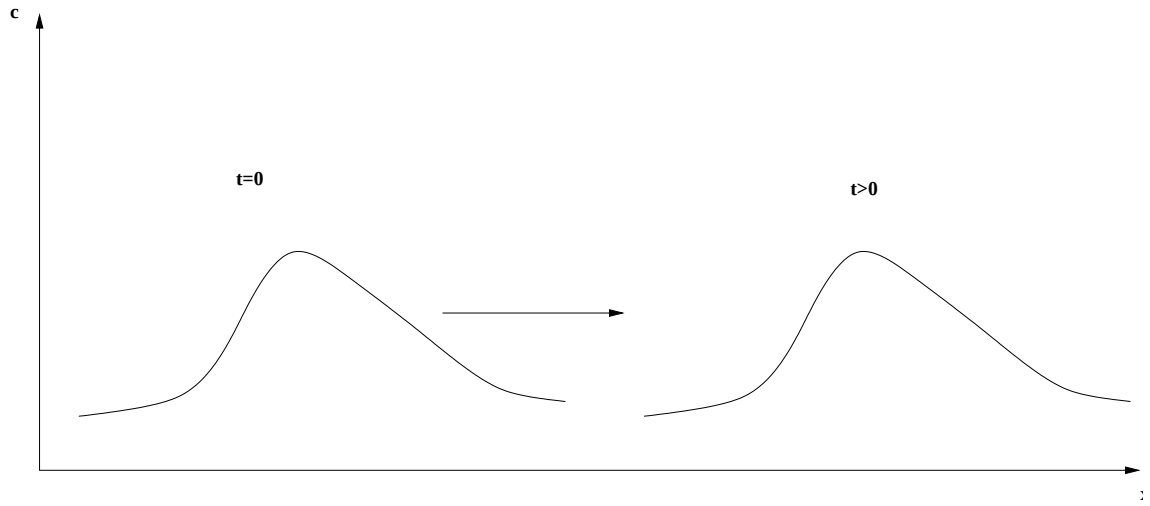

FIG. 4.1 – Advection d'un traceur

Pour préparer le cas général, on peut retrouver le résultat précédant en recourant à une approche lagrangienne (la méthode des caractéristiques). Soit $X_{x_0}(t)$ la courbe caractéristique issue de la position initiale x_0 :

$$\frac{dX_{x_0}}{dt} = V \ , \quad X_{x_0}(0) = x_0 \tag{4.3}$$

et soit $\bar{c}_{x_0}(t) = c(X_{x_0}(t), t)$ la concentration du traceur le long de cette courbe. On a évidemment :

$$\frac{d\bar{c}_{x_0}}{dt} = \frac{\partial c}{\partial t} + \frac{dX_{x_0}}{dt}\frac{\partial c}{\partial x} = 0 \tag{4.4}$$

c'est à dire que $\bar{c}_{x_0}(t)$ est constant et vaut donc $c_0(x_0)$. Ceci implique également que les courbes caractéristiques sont des droites :

$$X_{x_0}(t) = x_0 + Vt \tag{4.5}$$

On a donc la solution en (x, t) lorsque la vitesse est constante : les caractéristiques sont alors des droites parallèles et il ne passe qu'une caractéristique en (x, t), celle associée au point $x_0 = x - Vt$. Par conservation du traceur le long de la caractéristique, on a directement le résultat (figure 4.2).

L'extension du modèle précédent au cas d'un champ de vitesse à divergence nulle ($\operatorname{div} V = 0$) est immédiate en remarquant que l'on a alors :

$$\operatorname{div}(Vc) = V \cdot \nabla c \tag{4.6}$$

Dans le cas plus physique où le champ de vitesse V est relié à la densité du fluide "porteur" ρ par l'équation de continuité :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(V\rho) = 0 \tag{4.7}$$

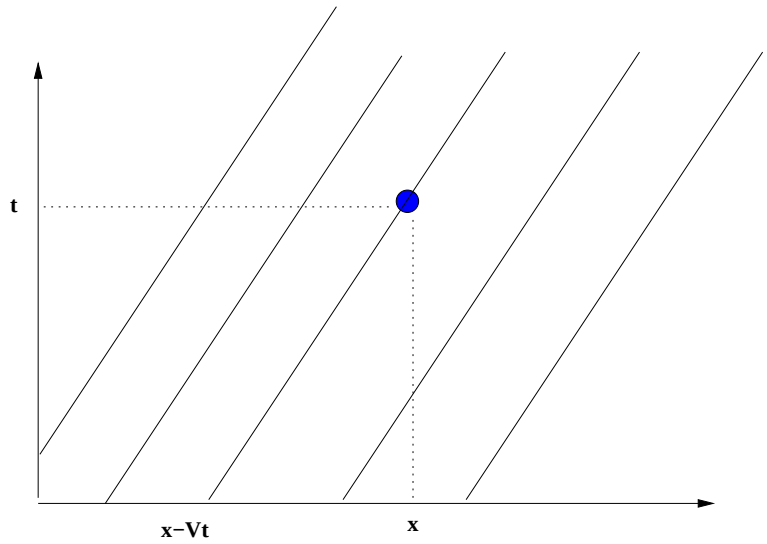

FIG. 4.2 – Méthode des caractéristiques. Cas d'une vitesse constante.

il est plus aisé de travailler pour décrire le traceur avec sa fraction massique que l'on notera par exemple $m = c/\rho$. Il est direct de constater que l'on a alors :

$$\frac{\partial m}{\partial t} + V \cdot \nabla m = 0 \tag{4.8}$$

et le long des caractéristiques, $m = c/\rho$ est conservé. Les caractéristiques correspondent alors exactement aux lignes de champ (tangentes au champ de vitesse et qui ne se coupent pas).

4.1.2 Propriétés qualitatives

Plusieurs propriétés découlent de manière directe de ce résultat :

1. la *positivité* des solutions si on part d'une condition initiale positive ;
2. la *monotonie* : il y a un "principe du maximum" pour l'advection, ie on ne doit pas créer d'extrema dans la solution qui n'existent pas dans la condition initiale ; il est à noter que dans le cas linéaire, "positivité" et "monotonie" sont étroitement associées.

Un schéma numérique va donc devoir respecter ces propriétés qualitatives clés.

4.1.3 Quelques schémas "évidents" de discrétisation spatiale

On se donne une discrétisation de l'axe des x selon (x_i) avec un pas de maillage supposé constant Δx (figure 4.3). On discrétise également le temps selon une suite (t_n) avec un pas de temps supposé egalement constant Δt.

On va dans un premier temps construire des schémas numériques avec la *méthode des lignes*, ie en discrétisant d'abord le problème en espace avant de résoudre le problème en

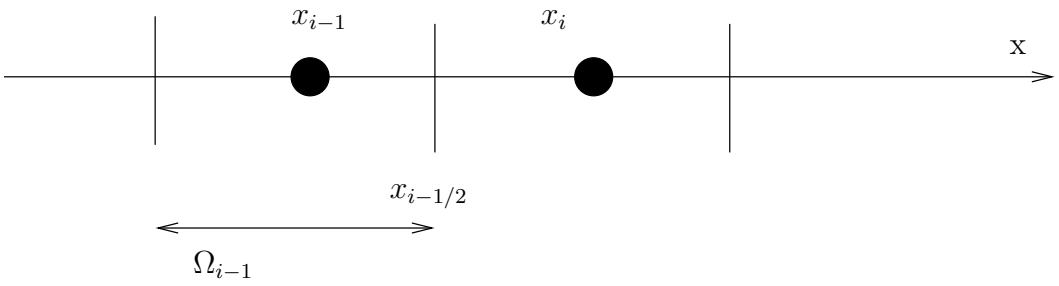

FIG. 4.3 – Maillage

temps. On notera par abus $c_i(t) \simeq c(x_i, t)$, l'indice i correspondant ici à un point de maille et non à une espèce (on advecte en parallèle les diverses espèces).

Une forme *conservative* du schéma numérique est définie par :

$$\frac{dc_i}{dt} = \frac{f_{i-1/2} - f_{i+1/2}}{\Delta x} \tag{4.9}$$

où le flux (entrant) $f_{i-1/2}$ approche le flux au niveau de la facette $x_{i-1/2}$ ($Vc(x - \Delta x/2, t)$). Cette forme garantit bien évidemment la conservation de la masse.

Un premier schéma évident, en supposant que $V > 0$, est de faire l'approximation pour le flux :

$$f_{i-1/2} = Vc_{i-1} \tag{4.10}$$

ou, ce qui revient au même, d'approcher par différences finies le gradient spatial selon :

$$c(x - \Delta x) = c(x) - \Delta x \frac{dc}{dx}(x) + \dots \tag{4.11}$$

ce qui conduit au schéma :

$$\frac{dc_i}{dt} = V \frac{c_{i-1} - c_i}{\Delta x} \tag{4.12}$$

On appelle classiquement ce flux le flux *upwind* ou *décentré* ou encore *donor-cell*, chacune des terminologies étant évidente.

Une seconde approche pourrait être d'approcher le flux entrant au niveau de la facette en $x_{i-1/2}$ par :

$$f_{i-1/2} = V \frac{c_{i-1} + c_i}{2} \tag{4.13}$$

ou de manière équivalente d'utiliser le développement de Taylor :

$$\frac{dc}{dx}(x) \simeq \frac{c(x + \Delta x) - c(x - \Delta x)}{2\Delta x} + \dots \tag{4.14}$$

On obtient alors le *flux centré* :

$$\frac{dc_i}{dt} = V \frac{c_{i-1} - c_{i+1}}{2\Delta x} \tag{4.15}$$

4.1.4 Discrétisation temporelle

On notera c_i^n la solution au temps t_n et en x_i. Par exemple, pour le flux upwind défini précédemment, plusieurs discrétisations en temps sont possibles, parmi lesquelles le schéma upwind explicite obtenu avec un schéma d'Euler explicite :

$$\frac{c_i^{n+1} - c_i^n}{\Delta t} = V \frac{c_{i-1}^n - c_i^n}{\Delta x} \tag{4.16}$$

Cette manière de procéder (discrétisation spatiale pour obtenir une EDO puis discrétisation en temps) définit ce que l'on appelle communément la *méthode des lignes*.

Notons que l'on aurait pu directement proposer ce schéma de discrétisation en discrétisant de manière conjointe temps et espace (on parle alors souvent de schéma DST pour "Direct Space Time"). Une manière directe de procéder aurait été d'utiliser la relation exacte :

$$c(x_i, t_{n+1}) = c(x_i - V\Delta t, t_n) \tag{4.17}$$

puis de chercher une interpolation par rapport aux points de discrétisation permettant d'estimer $c(x_i - V\Delta t, t_n)$. Dans le cas du schéma upwind, si $V > 0$ et $V\Delta t \leq \Delta x$ (on reviendra sur cette condition par la suite), il est licite d'interpoler $x_i - V\Delta t$ entre x_{i-1} et x_i. Une interpolation linéaire donne directement le schéma upwind écrit sous la forme :

$$c_i^{n+1} = (1-a)c_i^n + ac_{i-1}^n \tag{4.18}$$

avec a le nombre de Courant-Friedrichs-Lewy (CFL) défini dans la section suivante comme :

$$a = V\frac{\Delta t}{\Delta x} \tag{4.19}$$

De la même manière, on peut définir une méthode d'ordre 3 (le schéma DST3) par :

$$c_i^{n+1} = \gamma_{-2}c_{i-2}^n + \gamma_{-1}c_{i-1}^n + \gamma_0 c_i^n + \gamma_1 c_{i+1}^n \tag{4.20}$$

avec $\gamma_{-2} = -a(1-a^2)/6$, $\gamma_{-1} = a(2-a)(1+a)/2$, $\gamma_0 = (2-a)(1-a^2)/2$ et $\gamma_1 = -a(2-a)(1-a)/6$.

4.1.5 Stabilité et ordre d'un schéma

De manière identique à ce qui a été présenté pour les EDOs (chapitre 3), l'erreur de discrétisation numérique comprend deux éléments ;

- une erreur locale liée à l'erreur de discrétisation sur un pas de temps (fonction de l'*ordre* du schéma) ;
- une propagation des erreurs antérieures (dont le comportement est donné par une étude de *stabilité*).

L'erreur locale se calcule en remplaçant dans l'algorithme itératif la solution approchée c_i^n par la solution exacte $c(x_i, t_n)$. Par exemple, pour le schéma upwind, on définit le résidu ρ_i^n par :

$$\frac{c(x_i, t_{n+1}) - c(x_i, t_n)}{\Delta t} - V\frac{c(x_{i-1}, t_n) - c(x_i, t_n)}{\Delta x} = \rho_i^n \tag{4.21}$$

Par développement limité, on obtient aisément $\rho_i^n = O(\Delta t) + O(\Delta x)$.

On a alors directement pour l'erreur globale $\varepsilon_i^n = c(x_i, t_n) - c_i^n$ par soustraction :

$$\frac{\varepsilon_i^{n+1} - \varepsilon_i^n}{\Delta t} - V\frac{\varepsilon_{i-1}^n - \varepsilon_i^n}{\Delta x} = \rho_i^n \tag{4.22}$$

qui fait clairement apparaître deux contributions à l'erreur.

L'étude de la stabilité est liée à la propagation des erreurs au cours du temps. La stabilité du schéma numérique est classiquement étudiée par l'analyse dite de Neumann en considérant une condition initiale donnée par un mode de Fourier. En se restreignant à l'intervalle $[0, 2\pi]$ avec des conditions aux limites périodiques (sans perte de généralité), on considère alors le mode de Fourier $\exp(jkx)$ avec $j^2 = -1$ et on cherche une solution de la forme :

$$c_i^n = (r_k)^n \exp(jkx_i) \tag{4.23}$$

Le point clé est le comportement du coefficient d'amplification r_k, qui doit rester de norme inférieur à 1 pour tout mode de Fourier k. En effet, par superposition, pour toute condition initiale périodique $c_0(x) = \sum_k (c_0)_k \exp(jkx)$, on alors :

$$c_n^i = \sum_k (r_k)^n \exp(jkx_i) \tag{4.24}$$

soit :

$$||c_n||^2 = \sum_i |c_n^i|^2 \leq ||c_0||^2 \tag{4.25}$$

Par exemple, pour le schéma upwind explicite, on obtient directement :

$$\frac{c_i^{n+1}}{c_i^n} = 1 + a(\exp(jk\Delta x) - 1) \stackrel{\Delta}{=} r_k \tag{4.26}$$

La stabilité est assurée par $|r_k| \leq 1$ et ce pour tous les modes de Fourier. Autrement dit avec :

$$|r_k|^2 = 1 - 2a(1-a)(1 - \cos(k\Delta x)) \tag{4.27}$$

on obtient (on se place depuis le début dans le cas $V \geq 0$) :

$$a = \frac{V\Delta t}{\Delta x} \leq 1 \tag{4.28}$$

que l'on appelle classiquement la *condition de Courant-Friedrich-Lewy* (condition CFL).

Notons que cette condition aurait pu être obtenue en demandant à respecter la positivité de la solution écrite sous la forme :

$$c_i^{n+1} = (1-a)c_i^n + ac_{i-1}^n \tag{4.29}$$

Il est à noter que cette condition est d'autant plus pénalisante pour le choix des pas de temps que le champ de vitesse est élevé et surtout que le maillage est fin. Ceci signifie en pratique, par exemple dans le cas de la simulation atmosphérique, que la condition CFL est contraignante pour les applications à petite échelle.

4.1.6 Comportement qualitatif : notion d'EDP équivalente

Il est "physiquement" clair qu'une propriété qualitative importante des schémas numériques présentés ci-dessus va être la génération d'une *diffusion numérique* artificiellement créée par la discrétisation.

Par exemple, pour le schéma upwind, prendre pour le flux $f_{i-1/2} = Vc_{i-1}$ revient à considérer que toute la matière dans la cellule Ω_{i-1} (y compris celle qui vient de rentrer dans cette cellule) contribue immédiatement au flux sortant en $x_{i-1/2}$.

Il est à noter, toujours dans le même ordre d'idée qualitative, que le comportement du schéma centré sera probablement moins diffusif. En effet, le flux se calcule à partir du flux upwind en rajoutant une correction "antidiffusive" selon :

$$f_{i-1/2} = Vc_{i-1} + \frac{V}{2}(c_i - c_{i-1}) \tag{4.30}$$

Pour la terminologie, il suffit de remarquer que si $c_i > c_{i-1}$, la correction consiste à rajouter de la matière dans la cellule Ω_i et à enlever dans la cellule Ω_{i-1} (ce qui est l'inverse de ce qu'aurait produit un flux de diffusion).

Ces remarques intuitives sont confirmées par les résultats de la figure 4.4.

Le comportement qualitatif des schémas numériques peut s'étudier de manière plus systématique par le recours à la notion d'*EDP équivalente*, une EDP qu'approche le schéma numérique de manière plus fine que le modèle d'advection. L'étude qualitative du schéma se fait alors sur cette EDP, étant plus aisée sur un cas continu que sur un cas discret (voir ci-après pour s'en convaincre).

Par exemple, pour le schéma upwind, en allant plus loin dans le développement de Taylor :

$$\frac{c(x-\Delta x) - c(x)}{\Delta x} = -\frac{dc}{dx}(x) + \frac{\Delta x}{2}\frac{d^2c}{dx^2}(x) + O(\Delta x^2) \tag{4.31}$$

ce qui montre que le schéma upwind est en réalité une approximation à l'ordre 2 de l'EDP :

$$\frac{\partial c}{\partial t} + V\frac{\partial c}{\partial x} = \frac{V\Delta x}{2}\frac{\partial^2 c}{\partial x^2} \tag{4.32}$$

ce qui justifie évidemment le caractère *diffusif* du schéma.

Pour le schéma centré, avec :

$$\frac{c(x-\Delta x) - c(x+\Delta x)}{2\Delta x} = -\frac{dc}{dx}(x) - \frac{\Delta x^2}{6}\frac{d^3c}{dx^3}(x) + O(\Delta x^4) \tag{4.33}$$

l'EDP équivalente est (à l'ordre 4) :

$$\frac{\partial c}{\partial t} + V\frac{\partial c}{\partial x} = -\frac{V\Delta x^2}{6}\frac{\partial^3 c}{\partial x^3} \tag{4.34}$$

qui a un comportement *dispersif* (les modes de Fourier ne sont pas advectés à la même vitesse).

De manière générale, l'étude qualitative de l'EDP

$$\frac{\partial c}{\partial t} + V\frac{\partial c}{\partial x} = \alpha\frac{\partial^2 c}{\partial x^2} + \beta\frac{\partial^3 c}{\partial x^3}\ , \quad c_k(x,0) = A_k(0)\ \exp(jkx) \tag{4.35}$$

est directement donnée en cherchant une solution de la forme $A_k(t)\exp(jk(x-\omega t))$. En identifiant les parties réelles et imaginaires, on trouve :

$$c_k(x,t) = A_k(0)\exp(-k^2\alpha t)\exp(jk(x-Vt) - jk\beta t) \tag{4.36}$$

à comparer à la solution de l'équation d'advection $A_k(0)\exp(jk(x-Vt))$. On a donc bien un comportement diffusif lié à α et un comportement dispersif (un déphasage) lié à β.

Les solutions des méthodes "upwind", "centré" et "DST3" sont tracées dans la figure 4.4 pour une nombre de CFL égal à 0.4 après 20 pas de temps. On note le caractère très diffusif du schéma upwind, le caractère dispersif du schéma centré et le bon comportement du schéma DST3, qui apparaît moins diffusif que le schéma "upwind" mais pour lequel la positivité n'est malheureusement plus assurée.

4.1.7 Méthodes à limiteurs de flux

En revenant sur les propriétés qualitatives que l'on souhaitait voir satisfaites par le schéma numérique, il apparaît donc difficile de pouvoir concilier à la fois :

1. la positivité de la solution (de manière équivalente pour le cas linéaire, la monotonie, ie la non-création d'extrema artificiels) ;
2. une faible diffusion numérique.

Ces deux points sont évidemment clés si l'on cherche à suivre un traceur issu d'une émission ponctuelle accidentelle (par exemple un radioélément) dans le cas par exemple de la dispersion atmosphérique. Le critère de positivité est clair et il est à noter que l'on se trouve dans le cas le plus défavorable d'un *pulse* pour la condition initiale...

Réanalysons qualitativement les schémas upwind et centré présentés précédemment. Le cas général est celui des schémas à 3 points avec :

$$c_i^{n+1} = \alpha_i\, c_{i-1}^n + \beta_i\, c_i^n + \gamma_i\, c_{i+1}^n \tag{4.37}$$

1. Le critère de positivité implique que les coefficients α_i, β_i et γ_i sont positifs ou nuls. Notons que cette condition est très restrictive car elle permet de garantir la positivité dans tous les cas (ie dans le cas extrême où toutes les concentrations sont nulles sauf dans une maille à t_n). Dans le cas d'une situation "réelle", une condition de positivité plus souple (mais dépendant de la situation) pourrait être suffisante. C'est ce que l'on va utiliser plus loin pour définir les limiteurs de flux.
2. Le critère de conservation de masse implique que :

$$\alpha_{i+1} + \beta_i + \gamma_{i-1} = 1 \tag{4.38}$$

c'est à dire que deux jeux de coefficients sont suffisants pour définir le schéma.

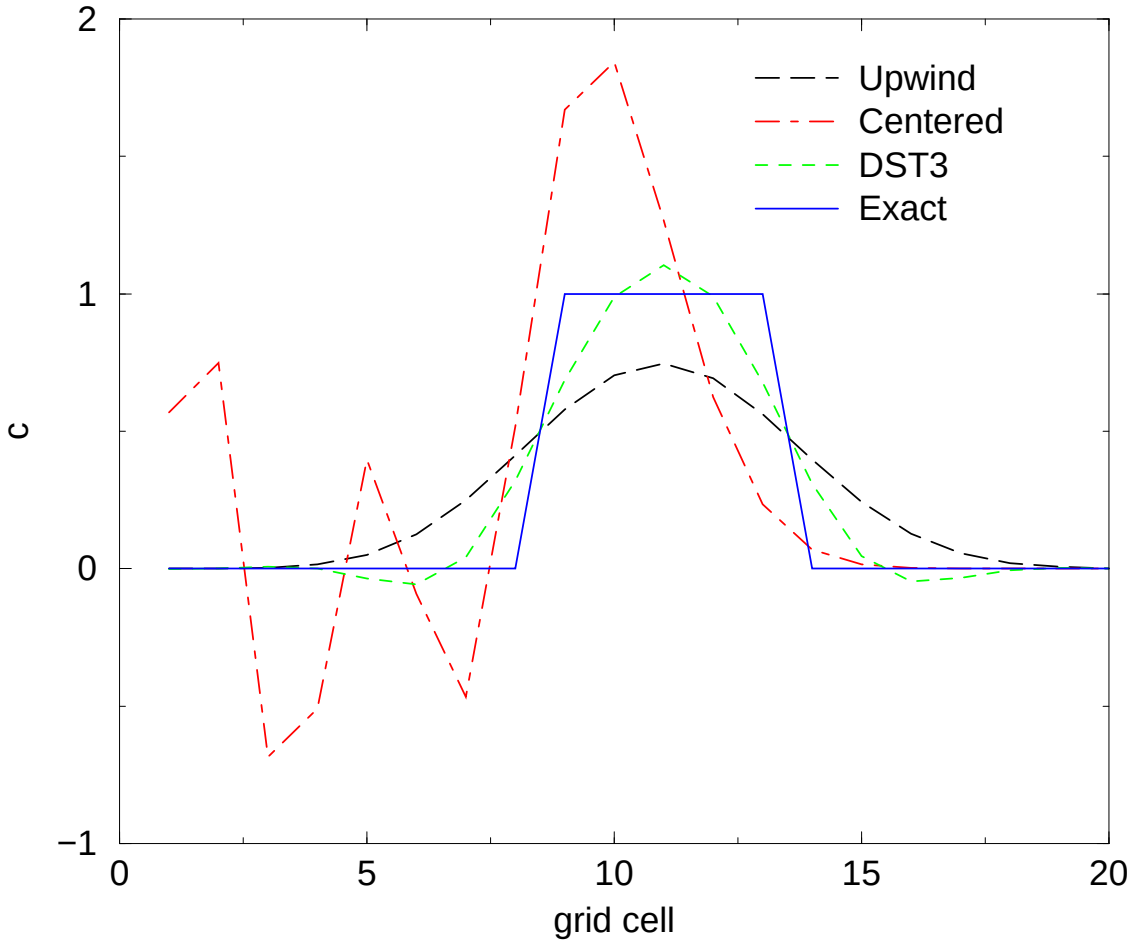

FIG. 4.4 – Comparaison des méthodes upwind, centrée et DST3 pour un nombre de CFL égal à 0.4 au bout de 20 itérations.

3. Afin d'étudier le comportement diffusif du schéma, on va le réécrire sous la forme équivalente suivante (on a noté $c_{i+1/2} = (c_i + c_{i+1})/2$) :

$$c_i^{n+1} = c_i^n - (a_{i+1/2}c_{i+1/2}^n - a_{i-1/2}c_{i-1/2}^n) + (\nu_{i+1/2}(c_{i+1}^n - c_i^n) - \nu_{i-1/2}(c_i^n - c_{i-1}^n)) \quad (4.39)$$

qui fait apparaître des termes de type advectif ($a_{i+1/2}$) et des termes de type diffusif ($\nu_{i+1/2}$). Les deux jeux de coefficients $a_{i+1/2}$ et $\nu_{i+1/2}$ sont donnés par :

$$\begin{aligned} \alpha_i &= \nu_{i-1/2} + \frac{1}{2}a_{i-1/2} \\ \gamma_i &= \nu_{i+1/2} - \frac{1}{2}a_{i+1/2} \\ \beta_i &= 1 - \nu_{i+1/2} - \frac{1}{2}a_{i+1/2} - \nu_{i-1/2} + \frac{1}{2}a_{i-1/2} \end{aligned} \quad (4.40)$$

La condition de positivité implique pour (α) et (γ) que :

$$\frac{1}{2}|a_{i+1/2}| \leq \nu_{i+1/2} \quad (4.41)$$

Par ailleurs, pour assurer l'ordre du schéma, $a = V\Delta x/\Delta t$ (ce qui justifie la notation). **Ceci signifie donc que la diffusion numérique ($\nu \geq 0$) est nécessaire pour garantir la positivité systématique**.

Si on revient au schéma upwind, on a directement avec $\gamma_i = 0$, $\nu = a/2$, c'est à dire que le schéma upwind est celui qui garantit, dans tous les cas de figure, la positivité avec une diffusion minimale... dont on a vu pourtant qu'elle n'est guère acceptable en pratique !

L'idée des méthodes à flux corrigés (ou des méthodes à limiteurs de flux) est de pouvoir descendre en dessous de la limite $\nu = a/2$ en fonction de la forme de la solution, puisqu'en pratique la condition de positivité est trop "frileuse". La stratégie va alors être d'utiliser le schéma upwind lorsque l'on est loin des gradients (son caractère diffusif n'est pas gênant et il garantit la positivité) mais d'utiliser un schéma d'ordre élevé peu diffusif proche des gradients.

Supposons que l'on dispose d'un schéma d'ordre élevé moins diffusif que le schéma upwind, que l'on écrit sous la forme relativement générale (pour son flux) suivante :

$$f_{i+1/2} = V(c_i + \Psi(\theta_i)(c_{i+1} - c_i)) \ , \quad \theta_i = \frac{c_i - c_{i-1}}{c_{i+1} - c_i} \quad (4.42)$$

θ_i est un indicateur du caractère "lisse" de la solution. Pour une solution constante, $\theta \simeq 1$, proche d'un gradient $\theta \ll 1$ ou $\gg 1$.

Par exemple, pour le schéma DST3, la fonction de flux s'écrit :

$$\Psi(\theta) = \frac{1}{6}[(2-a)(1-a) + (1-a)^2\theta] \quad (4.43)$$

Dans le cas général, on obtient après calcul la formule de récurrence suivante :

$$c_i^{n+1} = (1 - a\phi_i)c_i^n + a\phi_i c_{i-1} \ , \quad \phi_i = 1 + \frac{\Psi(\theta_i)}{\theta_i} - \Psi(\theta_{i-1}) \quad (4.44)$$

On va alors construire un schéma garantissant la positivité en imposant que $0 \leq a\phi_i \leq 1$: pour cela, on va remplacer Ψ par le flux limité Ψ_L avec :

$$0 \leq \Psi_L \leq 1 \ , \quad 0 \leq \Psi_L/\theta \leq \mu \tag{4.45}$$

où μ est un paramètre numérique à choisir. Ce choix de limitation des flux correspond au *limiteur de Koren-Sweby* (mais d'autres choix sont possibles).

Il suffit pour cela de choisir :

$$\Psi_L = \max(0, \min(1, \mu\theta, \Psi)) \tag{4.46}$$

Ceci donne comme condition de positivité $a(1+\mu) \leq 1$ et détermine μ en fonction du nombre de CFL a (localement). En pratique, on prend $a = (1-\nu)/\nu$.

4.1.8 Extension aux cas 2D et 3D

En pratique, l'advection a lieu dans un cas bi ou tridimensionnel et non pas dans une seule direction d'espace. Plusieurs approches sont alors possibles pour l'extension des méthodes précédentes :

- on peut effectuer un splitting directionnel, en résolvant de manière successive chacune des directions. Un inconvénient immédiat est alors la perte de monotonie, en particulier le fait qu'un champ de conditions initiales constant ne reste pas constant dans un champ de vitesse à divergence nulle.
  Des corrections sont alors nécessaires mais on ne précise pas plus ici.
- une alternative est d'effectuer une résolution couplée des directions en agrégeant les termes liés à chaque direction. Un inconvénient est cependant la sévérisation de la contrainte CFL. Par exemple pour un cas bidimensionnel (x, y) avec un champ de vitesse $V = (u, v)$, on obtient :

  $$\left(\frac{|u|}{\Delta x} + \frac{|v|}{\Delta y}\right) \Delta t \leq 1 \tag{4.47}$$

  évidemment plus contraignante que la contrainte unidimensionnelle.

4.2 Diffusion

4.2.1 Modèle

On considère ici la diffusion d'un traceur, de concentration c, sous l'effet d'une diffusion K :

$$\frac{\partial^2 c}{\partial t^2} = \mathrm{div}(K\nabla c) \ , \quad c(x, t=0) = c_0(x) \tag{4.48}$$

avec des conditions aux limites si nécessaire (domaine non borné).

4.2.2 Algorithme aux différences finies

La discrétisation numérique ne pose pas de problème numérique spécifique. On utilise classiquement une discrétisation à 3 points fondée sur un développement de Taylor. Par exemple, pour la méthode des lignes, la discrétisation spatiale donne :

$$\frac{dc_i}{dt} = \frac{K_{i+1/2}\dfrac{c_{i+1}-c_i}{\Delta x} - K_{i-1/2}\dfrac{c_i - c_{i-1}}{\Delta x}}{\Delta x} \tag{4.49}$$

avec $K_{i+1/2} = K(x_{i+1/2})$.

Cette EDO doit être résolue. Pour le cas d'une diffusion constante, si on utilise la méthode d'Euler explicite, on a directement :

$$\frac{c_i^{n+1} - c_i^n}{\Delta t} = \frac{K}{\Delta x^2}(c_{i+1}^n - 2c_i^n + c_{i-1}^n) \tag{4.50}$$

L'étude de stabilité de la discrétisation temporelle de cette approximation par différences finies peut se faire de manière classique par une analyse de Neumann en calculant les coefficients d'amplification associés aux modes de Fourier.

En utilisant la condition initiale $\exp(jkx_i)$ correspondant au mode k de Fourier, on a après un calcul aisé :

$$\frac{c_i^{n+1}}{c_i^n} = 1 - \frac{4\Delta t}{\Delta x^2}\sin^2(k\Delta x/2) \triangleq r_k \tag{4.51}$$

La stabilité est assurée par $|r_k| \leq 1$ et ce pour tous les modes de Fourier. Autrement dit :

$$\frac{K\Delta t}{\Delta x^2} \leq \frac{1}{2} \tag{4.52}$$

que l'on appelle classiquement la *condition de Fourier*. Cette condition de stabilité n'est évidemment pas nécessaire pour des résolutions implicites, qui sont en général préférées pour l'intégration de la diffusion. Des algorithmes d'inversion de matrices doivent alors être spécifiés. Du fait de la forme tridiagonale de la matrice, l'inversion se fait aisément (algorithme de Thomas par exemple).

4.3 Exercices

4.3.1 Discrétisation de la diffusion

Écrire le schéma de discrétisation centré à trois points pour le terme de diffusion, dans le cas unidimensionnel et pour un coefficient de diffusion constant.

On considère l'équation de diffusion sur l'intervalle $[0, 1]$ divisé en n sous-intervalles. On note $h = \dfrac{1}{n}$ le pas de discrétisation et c la solution de l'équation discrétisée. On impose des conditions de Neumann aux frontières, approchées par des discrétisations d'ordre 1 : $\dfrac{c_1 - c_0}{h} = 0$ et $\dfrac{c_n - c_{n-1}}{h} = 0$. On note $\tilde{c} = (c_1, c_2, \cdots, c_{n-1})$. *Écrire l'équation de diffusion discrétisée (spatialement), par le schéma précédent, sous la forme* $\dfrac{\mathrm{d}\tilde{c}}{\mathrm{d}t} = kA\tilde{c}$ *où A est une matrice à déterminer.*

4.3.2 Stabilité L^2

On considère, en dimension 1, l'équation d'advection linéaire $\dfrac{\partial c}{\partial t} + V\dfrac{\partial c}{\partial x} = 0$. La solution discrétisée est notée c_j^n où n est l'indice temporel et j l'indice spatial. On note Δt le pas temporel et Δx le pas temporel. Le schéma numérique est tel que :

$$c_j^{n+1} = \sum_{k=-p}^{q} a_k c_{k+j}^n$$

On s'assure de la stabilité $\mathcal{L}^2(\mathbb{Z})$ en introduisant $\hat{c}^n(z) = \sum_{j\in\mathbb{Z}} c_j^n e^{-ij\Delta x z}$, avec $i^2 = -1$. *Écrire $\hat{c}^{n+1}(z)$ sous la forme $R(z)\hat{c}^n(z)$.*

En déduire pourquoi on s'assure de la stabilité en s'assurant que $\|c^n\|_2$ est borné si $c_j^0 = e^{ij\Delta x z}$, ceci pour tout $z \in [-\pi, \pi]$.

Discrétiser l'équation d'avection par un schéma centré en espace et un schéma d'Euler explicite en temps. Étudier, par la méthode précédente, la stabilité du schéma obtenu.

Étudier la stabilité du schéma de Lax-Friedrichs :

$$c_j^{n+1} = \frac{c_{j+1}^n + c_{j-1}^n}{2} - \frac{\Delta t}{2\Delta x}\left(V c_{j+1}^n - V c_{j-1}^n\right)$$

4.3.3 Schéma de Lax-Wendroff

En utilisant la formule de Taylor à l'ordre 2, retrouver le schéma de Lax-Wendroff pour l'advection :

$$c_j^{n+1} = c_j^n - \frac{V\Delta t}{2\Delta x}(c_{j+1}^n - c_{j-1}^n) + \frac{V^2\Delta t^2}{2\Delta x^2}(c_{j+1}^n - 2c_j^n + c_{j-1}^n)$$

Quel est l'ordre du schéma ?

4.3.4 Variation totale décroissante

Montrer que le schéma "upwind" est à variation totale décroissante (VTD[1]), c'est-à-dire que : $\sum_j |c_j^{n+1} - c_{j-1}^{n+1}| \leq \sum_j |c_j^n - c_{j-1}^n|$.

[1]TVD en anglais : "total variation diminishing".

Bibliographie

[1] I. Ahmad and M. Berzins. An algorithm for odes from atmospheric dispersion problems. *Applied Numerical Mathematics*, 25 :137–149, 1997.

[2] R.C. Aiken. *Stiff computation.* Oxford University Press, 1985.

[3] J.F. Andrus. Numerical solution of systems of odes separated into subsystems. *SIAM J.Numer.Anal*, 16(4), 1979.

[4] G.D. Byrne and A.C. Hindmarsh. Stiff ode solvers : a review of current and coming attractions. *J.Comp.Phys.*, 70 :1–62, 1987.

[5] D. Cariolle. Modèle unidimensionnel de la chimie de l'ozone. *Planet Sp. Sc.*, 31(9), 1983.

[6] D.P. Chock, S.L. Winkler, and Pu Sun. Comparison of stiff chemistry solvers for air quality modeling. *Env.Sci.Tech.*, 28 :1882–1892, 1994.

[7] P.G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation.* Masson, 1990.

[8] Curtis and Hirshfelder. Integration of stiff equations. *Proceedings of the National Academy of Science*, 38 :235–243, 1952.

[9] D. Dabdub and J.H. Seinfeld. Extrapolation techniques used in the solution of stiff odes associated with chemical kinetics of air quality models. *Atm. Env.*, 29(3) :403–410, 1995.

[10] S. Elliott, R.P. Turco, and M.Z. Jacobson. Tests on combined projection/forward differencing integration for stiff photochemical chemical family systems at long time step. *Computers Chem.*, 17(1) :91–102, 1993.

[11] M.W. Gery, G.Z. Whitten, J.K. Killus, and M.C. Dodge. A photochemical kinetics mechanism for urban and regional scale computer modeling. *J. Geophys. Research*, 94(D10) :12925–12956, 1989.

[12] Gong and Cho. A numerical scheme for the integration of the gas-phase chemical rate equations in 3d atmospheric models. *Atmos.Environ.*, 27A :2591–2611, 1993.

[13] J. Graf and N. Moussiopoulos. Intercomparison of two models for the dispersion of chemically reacting pollutants. *Beitr.Phys.Atmosph.*, 64(1) :13–25, Febr. 1991.

[14] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic problems.* Springer, 1991.

[15] O. Hertel, R. Berkowicz, and J. Christensen. Test of two numerical schemes for use in atmospheric transport-chemistry models. *Atmos.Environ.*, 27A(16) :2591–2611, 1993.

[16] E. Hesstvedt, O. Hov, and I.S.A. Isaksen. Qssas in air polution modelling : comparison of two numerical schemes for oxydant prediction. *Int.J.Chem.Kinet.*, 10 :971–994, 1978.

[17] A.C. Hindmarsh. *Scientific computing*, chapter ODEPACK : a systematized collection of ODE solvers, pages 55–74. North Holland, 1983.

[18] W.H. Hundsdorfer. Numerical solution of advection-diffusion-reaction equations. Technical Report NM-N9603, CWI, 1996.

[19] W.H. Hundsdorfer and J.G. Verwer. A note on splitting errors for advection-reaction equations. Technical Report NM-R9424, CWI, 1994.

[20] M.Z. Jacobson and R.P. Turco. Smvgear : a sparse-matrix, vectorized gear code for atmospheric models. *Atm. Env.*, 28(2) :273–284, 1994.

[21] C.T. Kelley. *Iterative methods for linear and nonlinear equations*. SIAM, 1995.

[22] C. Kessler, A. Griesel, and J.G. Verwer. A rosenbrock solver in chemistry-transport modelling : what about the speed in 3d ? In *Proceedings EUROTRAC-2 Symposium*, 1998.

[23] O. Knoth and R. Wolke. *Air Pollution Modelling and its applications X*, chapter A comparison of fast chemical kinetic solvers in a simple vertical diffusion model. Plenum Press, NY, 1994.

[24] L. Lanser and J.G. Verwer. Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling. In *Proceedings 2nd Meeting on Numerical methods for differential equations*. Coimbra, Portugal, February 1998.

[25] B. Larrouturou. Modélisation mathématique et numérique pour les sciences de l'ingénieur. Cours Ecole Polytechnique, Majeure Sciences de l'Ingénieur et Calcul Scientifique, 1994.

[26] R.J. LeVeque and H.C. Yee. A study of numerical methods for hyperbolic conservation laws with stiff source terms. *J.Comp.Phys.*, (86) :187–210, 1990.

[27] G.I. Marchuk. *Mathematical models in environmental problems*, volume 16. North Holland, 1986.

[28] J.C. Miellou. Existence globale pour une classe de systèmes paraboliques semi-linéaires modélisant le problème de la stratosphère : la méthode de la fonction agrégée. *CRAS*, 299(14) :723, 1984.

[29] M.T. Odman, N. Kumar, and A.G. Russel. A comparison of fast chemical kinetic solvers for air quality modeling. *Atm. Env.*, 26A(9) :1783–1789, 1992.

[30] J.B. Perot. An analysis of fractional step method. *J.C.P.*, 108 :51–58, 1993.

[31] J.S. Rosenbaum. Conservation properties of numerical integration methods for systems of odes. *J.Comp.Phys.*, (20) :259–267, 1976.

[32] H.H. Rosenbrock. Some general implicit processes for the numerical solution of differential equations. *Computer j.*, 5 :329–330, 1963.

[33] A. Sandu, F.A. Potra, G.R. Carmichael, and V. Damian. Efficient implementation of fully implicit methods for atmospheric chemical kinetics. *J.Comp.Phys.*, 129 :101–110, 1996.

[34] A. Sandu, J.G. Verwer, J.G. Blom, E.J. Spee, and G.R. Carmichael. Benchmarking stiff odes solvers for atmospheric chemistry problems ii : Rosenbrock solvers. *Atm. Env.*, 31 :3459–3472, 1997.

[35] A. Sandu, J.G. Verwer, M. Van Loon, G. Carmichael, F.A. Potra, D. Dabdub, and J.H. Seinfeld. Benchmarking stiff odes solvers for atmospheric chemistry problems i : implicit versus explicit. *Atmos.Environ.*, 31 :3151–3166, 1997.

[36] J.M. Sanz-Serna. *The State of the art in numerical analysis*, chapter Geometric integration, pages 121–143. Clarendon Press, Oxford, 1997.

[37] R.D. Saylor and G.D. Ford. On the comparison of numerical methods for the integration of kinetic equations in atmospheric chemistry and transport models. *Atm. Env.*, 29(19) :2585–2593, 1995.

[38] L.F. Shampine. Solving odes in quasi steady state. In *Bienefeld Conference*, 1980.

[39] L.F. Shampine. Stiffness and the automatic selection of ode codes. *J.of Computational Physics*, 54, 1984.

[40] D. Shyan-Shu Shieh, Y. Chang, and G.R. Carmichael. The evaluation of numerical techniques for solution of stiff odes arising from chemical kinetics problems. *Env. Soft.*, 3(1), 1988.

[41] S. Skelboe and Z. Zlatev. *Numerical analysis and its applications*, chapter Exploiting the natural partitioning in the numerical solution of ODE systems arising in atmospheric chemistry, pages 458–465. Springer, 1997.

[42] B.P. Sommeijer, P.J. Van der Houwen, and J.G. Verwer. On the treatment of time-dependent boundary conditions in splitting methods for parabolic differential equations. *Int.J.for Num. Met. in Eng.*, 17 :335–346, 1981.

[43] E.J. Spee. Coupling advection and chemical kinetics in a global atmosperic test model. Technical Report NM R9508, CWI, 1995.

[44] E.J. Spee. *Numerical methods in global transport-chemistry models.* PhD thesis, Univ. Amsterdam, 1998.

[45] B. Sportisse. *Contribution à la modélisation des écoulements réactifs : réduction des modèles de cinétique chimique et simulation de la pollution atmosphérique.* PhD thesis, Ecole Polytechnique, April 1999.

[46] B. Sportisse. Assimilation de données et modélisation inverse. Cours ENSTA, 2005.

[47] B. Sportisse. Modélisation de la pollution atmosphérique. Cours ENPC, 2005.

[48] G. Strang. On the construction and comparison of difference schemes. *SIAM J.Numer.Anal.*, 5 :506–517, 1968.

[49] Pu Sun, D. Chock, and S.L. Winkler. An implicit-explicit hybrid solver for a system of stiff kinetic equations. *J. Comp. Physics*, 115 :515, 1994.

[50] M. Van Loon. *Numerical methods in smog prediction.* PhD thesis, Univ. Amsterdam, 1996.

[51] J.G. Verwer. Gauss-seidel iteration for stiff odes from chemical kinetics. *SIAM J.Sci.Comput.*, 15 :1243–1250, 1994.

[52] J.G. Verwer and J. Blom. On the coupled solution of diffusion and chemistry in air pollution models. In Akademie Verlag, editor, *Proceedings of ICIAM 95*, ZAMM, issue 4 : applied sciences, pages 454–457, 1996.

[53] J.G. Verwer, J. Blom, M. Van Loon, and E.J. Spee. A comparison of stiff odes solvers for atmospheric chemistry problems. *Atmos.Environ.*, 30 :49–58, 1996.

[54] J.G. Verwer, J.G. Blom, and W.H. Hundsdorfer. An implicit-explicit approach for atmospheric transport-chemistry problems. *Appl. Num. Math.*, 20 :191–209, 1996.

[55] J.G. Verwer and H.B. de Vries. Global extrapolation and first-order splitting method. *SIAM J. Sci.Stat.Comput.*, 6(3), 1985.

[56] J.G. Verwer, W.H. Hundsdorfer, and J.G. Blom. Numerical time integration for air pollution models. In *Proceedings of the Conference APMS'98*. ENPC-INRIA, October 26-29 1998.

[57] J.G. Verwer and D. Simpson. Explicit methods for stiff odes from atmospheric chemistry. *Appl. Num. Math.*, 18 :413–430, 1995.

[58] J.G. Verwer and M. Van Loon. An evaluation of explicit pseudo-steady state approximation schemes for stiff ode systems from chemical kinetics. *J.Comp.Phys.*, 113 :347–352, 1994.

[59] J.H. Verwer, E.J. Spee, J.G. Blom, and W.H. Hundsdorfer. A second order rosenbrock method applied to photochemical dispersion problem. *SIAM J. Sc. Comput.*, 20(4) :1456–1480, 1999.

[60] A.I. Vol'pert and S.I. Hudjaev. *Analysis in classes of discontinuous functions and equations of mathematical physics.*, chapter 12. Martinus Nijhoff, 1985.

[61] D.S. Watkins and R.W. Hansonsmith. The numerical solution of separably stiff systems by precise partitioning. *ACM Trans. on math. soft.*, 9(3), 1983.

[62] N.N. Yanenko. *The method of fractional steps*. New-York, 1971.

[63] T.R. Young and J.P. Boris. A numerical technique for solving stiff ode associated with the chemical kinetics of reaction flow problems. *J.Phys.Chem.*, 81 :2424, 1977.