

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'enseignement supérieur et la recherche scientifique

Université de batna 02

Faculté des Mathématiques et Informatique

Département des Mathématiques



Support de cours

Option : 3 ème année Licence Mathématique Académique

Méthodes Numériques pour les ODEs et EDP

Hanachi Adalet

Année Universitaire : 2019 - 2020



Table des matières

1	Introduction générale	5
1.1	L'analyse numérique des équations différentielles ordinaires et aux dérivées partielles	5
1.2	Les principales méthodes de discrétisation	6
1.2.1	Les méthodes de différences finies et volumes finis	6
1.2.2	Les méthodes variationnelles, méthodes d'éléments finis	6
1.2.3	Les méthodes spectrales	6
2	Rappels sur les résultats d'existence de la solution des EDO(s)	7
2.1	Introduction	7
2.2	Problème de Cauchy	8
2.2.1	Interprétation physique	8
2.2.2	Solutions de l'équation différentielle ordinaire	8
2.2.3	Cas de la dimension 1	9
2.2.4	Existence et unicité de la solution du problème de Cauchy	9
3	Résolution numérique du problème de Cauchy	11
3.1	Principe générale des méthodes numériques	11
3.1.1	Propriétés des méthodes numériques	11
3.1.2	Les principales méthodes numériques	12
3.2	Analyse des méthodes à un pas (convergence, stabilité et consistance)	15
3.2.1	Applications sur les méthodes numériques précédentes	17
3.3	Méthodes à pas multiples	19
3.3.1	Analyse de convergence des méthodes à q pas	21
3.3.2	Application sur les méthodes précédentes	23
4	Méthode de différences finies	25
4.1	Introduction	25
4.1.1	Principe de la méthode	25
4.2	Différence finies pour l'équation de la chaleur	25
4.3	Convergence du schéma numérique	29
4.3.1	Consistance	29
4.4	Stabilité	30
4.4.1	Stabilité en norme L^∞	31
4.4.2	Stabilité en norme L^2	32
4.5	Schémas multiniveaux	33
4.6	Erreurs de troncature et stabilité de divers schémas pour l'équation de la chaleur	35

Chapitre 1

Introduction générale

1.1 L'analyse numérique des équations différentielles ordinaires et aux dérivées partielles

Pour faire l'implémentation numérique (à l'aide d'un outil informatique) des solutions d'un problème réel, on passe par les étapes suivantes :

1. **Description qualitative des phénomènes physiques.** Dans cette étape les spécialistes des phénomènes (ingénieurs, biologistes, chimistes, etc....) ont répertorié tous les mécanismes qui entrent en jeu dans le problème qu'on étudie.
2. **Modélisation.** Il s'agit à partir de la description qualitative précédente, d'écrire un modèle mathématique. On supposera ici que ce modèle amène à un système d'équations aux dérivées partielles (EDP). Selon les hypothèses effectuées, la modélisation peut aboutir à plusieurs modèles, plus ou moins complexes. Dans la plupart des cas, on ne saura pas calculer une solution analytique, explicite, du modèle ; on devra faire appel à des techniques de résolution approchée.
3. **Analyse mathématique du modèle mathématique.** Même si l'on ne sait pas trouver une solution explicite du modèle, il est important d'étudier les propriétés mathématiques, dans la mesure du possible. Il est bon de se poser les questions suivantes : Le problème est-il bien posé ? c'est-à-dire y-a-t'il existence et unicité de la solution ? Les propriétés physiques auxquelles on s'attend sont-elles satisfaites par les solutions du modèle mathématique ? Si l'inconnue est une concentration, par exemple, peut-on prouver que la solution du modèle sensé la représenter est toujours positive ? Y a-t-il continuité de la solution par rapport aux données ?
4. **Discrétisation et résolution numérique.** Un problème posé sur un domaine continu (espace - temps) n'est pas résoluble tel quel par un ordinateur, qui ne peut traiter qu'un nombre fini d'inconnues. Pour se ramener à un problème en dimension finie, on discrétise l'espace et ou le temps. Si le problème original est linéaire on obtient un système linéaire. Si le problème original est non linéaire (par exemple s'il s'agit de la minimisation d'une fonction) on aura un système non linéaire à résoudre par une méthode de Newton..
5. **Analyse numérique.** Il s'agit maintenant de l'analyse mathématique du schéma numérique. En effet, une fois le problème discret obtenu, il est raisonnable de se demander si la solution de ce problème est proche, et en quel sens, du problème continu. De même, si on doit mettre en œuvre une méthode itérative pour le traitement, il faut étudier la convergence de la méthode itérative proposée.

6. **Mise en œuvre, programmation et analyse des résultats.** La partie mise en œuvre est une grosse consommatrice de temps. Actuellement, de nombreux codes commerciaux existent, qui permettent en théorie de résoudre "tous" les problèmes. Il faut cependant procéder à une analyse critique des résultats obtenus par ces codes, qui ne sont pas toujours compatibles avec les propriétés physiques attendues...

1.2 Les principales méthodes de discrétisation

1.2.1 Les méthodes de différences finies et volumes finis

On considère un domaine physique $\Omega \subset \mathbb{R}^n$, où n est la dimension de l'espace. Le principe des méthodes de différences finies (DF) consiste à se donner un certain nombre de points du domaine, qu'on notera $\{x_1, \dots, x_N\} \subset \Omega$. On approche l'opérateur différentiel en espace en chacun des x_i par des quotients différentiels.

Les méthodes de volumes finis (VF) sont adaptées aux équations de conservation et utilisées en mécanique des fluides depuis plusieurs décennies. Le principe consiste à découper le domaine Ω en des "volumes de contrôle" on intègre ensuite l'équation de conservation sur les volumes de contrôle.

1.2.2 Les méthodes variationnelles, méthodes d'éléments finis

On met le problème d'équations aux dérivées partielles sous la forme suivante, dite variationnelle :

$$\begin{cases} a(u, v) = (f, v)_H, & \forall v \in H \\ u \in H \end{cases}$$

où H est un espace de Hilbert bien choisi (par exemple parce qu'il y a existence et unicité de la solution dans cet espace), $(\cdot, \cdot)_H$ le produit scalaire sur H et a une forme bilinéaire sur H . Dans un tel cadre fonctionnel, la discrétisation consiste à remplacer H par un sous espace de dimension finie H_k , construit par exemple à l'aide de fonctions de base éléments finis.

$$\begin{cases} a(u_k, v_k) = (f, v_k)_{H_k}, & \forall v_k \in H_k \\ u_k \in H_k \end{cases} .$$

1.2.3 Les méthodes spectrales

L'idée de ces méthodes est de chercher une solution approchée sous forme d'un développement sur une certaine famille de fonctions. On peut par exemple écrire la solution approchée sous la forme : $u = \sum_{i=1}^n \alpha_i(u) p_i$ où les p_i sont des fonctions polynomiales. On choisit la base p_i de manière à ce que les dérivées de α_i et p_i soient faciles à calculer. Ces dernières méthodes sont réputées coûteuses, mais précises. Elles sont d'ailleurs plus souvent utilisées comme aide à la compréhension des phénomènes physiques sur des problèmes modèles que dans pour des applications industrielles.

Chapitre 2

Rappels sur les résultats d'existence de la solution des EDO(s)

2.1 Introduction

L'objectif de ce chapitre est de donner les résultats d'existence des solutions pour les équations différentielles ordinaires.

Définition 2.1.1. Soient I un ouvert de \mathbb{R} et U un ouvert de \mathbb{R}^n . Soit

$$f : I \times U \rightarrow \mathbb{R}^n$$

une application continue. On appelle équation différentielle ordinaire la relation :

$$y' = f(t, y) \quad (t, y) \in I \times U \quad (2.1)$$

Exemple 2.1.1. 1. $y' = a(t)y + b(t)$.

2. $ty' - y = f(y)$

3. l'équation de Bernoulli $y' + a(t)y + b(t)y^\alpha = 0$ avec $\alpha \in \mathbb{R} \setminus \{0, 1\}$

4. l'équation de Riccati $y' + a(t)y + b(t)y^2 + c(t) = 0$

Remarque 2.1.1. 1 On peut écrire l'équation (E) sous forme système différentiel du premier ordre; on a $y = (y_1, y_2, \dots, y_n)$ et $f = (f_1, f_2, \dots, f_n)$ donc l'équation (2.1) apparaît comme un système différentiel à n inconnues y_1, y_2, \dots, y_n

$$(E) \begin{cases} y'_1 = f_1(t, y_1, y_2, \dots, y_n) \\ y'_2 = f_2(t, y_1, y_2, \dots, y_n) \\ \vdots \\ y'_n = f_n(t, y_1, y_2, \dots, y_n) \end{cases}$$

2 Si l'équation différentielle d'ordre supérieure

$$y^p = f(t, y, y' \dots y^{(p-1)})$$

on peut ramener l'équation à un système différentielle du premier ordre, en posant $z = (z_1, z_2, \dots, z_p)$ et $z_1 = y$, on obtient le système suivant

$$\begin{cases} z'_1 = z_2 \\ z'_2 = z_3 \\ \vdots \\ z'_{p-1} = z_p \\ z'_p = f(t, z_1, z_2, \dots, z_p) \end{cases}$$

Exemple 2.1.2. Soit l'équation différentielle du second ordre

$$y'' = 3y' - ty$$

Posons $z_1 = y$ et $z_1' = z_2$, donc l'équation devient

$$\begin{cases} z_1' = z_2 \\ z_2' = 3z_2 - tz_1 \end{cases} = \begin{pmatrix} 0 & 1 \\ -t & 3 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

Exercice 2.1.1. Trouver le système différentiel associé à l'équation de Vonder Pol

$$x''(t) + x(t) = \mu(1 - x^2(t))x'(t)$$

2.2 Problème de Cauchy

Le problème de Cauchy (problème aux valeurs initiales) consiste à chercher la solution d'une EDO scalaire ou vectorielle satisfaisant des conditions initiales

$$(PC) \begin{cases} y' = f(t, y) & (t, y) \in I \times U \\ y(t_0) = y_0 & (t_0, y_0) \in I \times U \end{cases} \quad (2.2)$$

Exemple 2.2.1.

$$\begin{cases} y'(t) = y(t) - 2\frac{t}{y(t)} \\ y(0) = 1 \end{cases}$$

2.2.1 Interprétation physique

Dans la plupart des situations, la variable t représente le temps et $y = (y_1, y_2, \dots, y_n)$ une famille de paramètres décrivant l'état d'un système matériel donné. L'équation (2.1) traduit physiquement la loi d'évolution du système considéré en fonction de temps et de la valeur des paramètres. Donc la résolution du problème de Cauchy revient à déterminer l'évolution du système au cours du temps sachant qu'en $t = t_0$ le système décrit par les paramètres $y_0 = (y_{01}, y_{02}, \dots, y_{0n})$ on dit que (t_0, y_0) est la donnée initiale du problème de Cauchy.

2.2.2 Solutions de l'équation différentielle ordinaire

Définition 2.2.1. Une solution de l'équation (2.1) sur un intervalle $I \subset \mathbb{R}$ est une fonction dérivable $y : I \rightarrow \mathbb{R}^n$ telle que

- 1 $\forall t \in I \quad (t, y(t)) \in I \times U$
- 2 $\forall t \in I \quad ; y' = f(t, y).$

Exemple 2.2.2. Soit l'équation

$$y' = f(t, y) = \frac{t}{y}$$

alors $f : \mathbb{R} \times \mathbb{R}^* \rightarrow \mathbb{R}$, montrons que $y(t) = t$ est une solution sur \mathbb{R}_+^*

- 1- on a $\forall t \in]0, \infty[\Rightarrow y(t) = t \in \mathbb{R}^*$.
- 2- $\forall t \in]0, \infty[\quad , y(t) = t$ est dérivable de plus elle vérifie $y' = 1 = \frac{t}{y}$

2.2.3 Cas de la dimension 1

L'équation (2.1) se réécrit

$$y' = \frac{dy}{dt} = f(t, y) \quad (t, y) \in I \times U \subset \mathbb{R} \times \mathbb{R}$$

Solution maximale et globale d'une équation du premier ordre

Définition 2.2.2. Soient $y : I \rightarrow \mathbb{R}$ et $\tilde{y} : \tilde{I} \rightarrow \mathbb{R}$ deux solutions de (2.1). On dit que \tilde{y} est prolongement de y si $\tilde{I} \supseteq I$ et $\tilde{y}|_I = y$.

Définition 2.2.3. On dit qu'une solution $y : I \rightarrow \mathbb{R}$ est maximale si y n'admet pas de prolongement \tilde{y} .

Théorème 2.2.1. Toute solution y se prolonge en une solution maximale \tilde{y} (pas nécessairement unique).

Définition 2.2.4. Une solution globale est solution définie sur l'intervalle tout entier.

Remarque 2.2.1. Toute solution globale est maximale, mais l'inverse est fausse.

Exemple 2.2.3. Soit l'équation $y' = -y^2$ sur \mathbb{R} . Cherchons les solutions de l'équation

1- On a $y(t) = 0$ est une solution de l'équation.

2- Si $y \neq 0$, donc on peut écrire l'équation sous la forme $\frac{y'}{y^2} = -1 \Leftrightarrow y(t) = \frac{1}{t+c}$. Donc on a définie deux solutions de l'équation définies respectivement sur $]-\infty, -c[$ et sur $]-c, \infty[$, ces solutions sont maximales mais non globales. La solution globale de l'équation est $y(t) = 0$.

2.2.4 Existence et unicité de la solution du problème de Cauchy

Définition 2.2.5. Soit f une application définie sur $I \times U \subset \mathbb{R}$. S'il existe une constante $L > 0$ indépendante de t, x et y telle que

$$|f(t, x) - f(t, y)| \leq L|x - y| \quad \forall x, y \in U, \forall t \in I$$

alors f est dite Lipschitzienne de rapport L sur $I \times U$ ou simplement L -Lipschitzienne.

Théorème 2.2.2. (Cauchy Lipschitz) Si f est une application définie sur $I \times U$ continue et L -Lipschitzienne par rapport à y , alors le problème de Cauchy (2.2) admet une solution unique sur I et ceci pour toute conditions initiale $(t_0, y_0) \in I \times U$

Proposition 2.2.1. Une condition suffisante pour que les hypothèses du théorème soient vérifiées est que f soit dérivable par rapport à y et que sa dérivée soit bornée.

Exemple 2.2.4. On considère le problème de Cauchy

$$(PC) \begin{cases} y' = y^2 = f(t, y) \\ y(0) = 1 \end{cases}$$

la fonction f est de classe C^1 , donc elle est Lipschitzienne sur les bornés (mais pas globalement Lipschitzienne). D'après le théorème de (2.2.2), il existe une solution unique maximale de l'équation $y' = y^2$

$$y' = y^2 \Leftrightarrow y(t) = \frac{1}{1-t}$$

on a $\lim_{t \rightarrow 1} y(t) = \infty$, donc le temps maximal d'existence est $T_M = 1$, la solution maximale est

$$y(t) = \frac{1}{1-t} \quad \text{sur } [0, 1[$$

Équivalence du problème de Cauchy avec la résolution d'une équation intégrale

Lemme 2.2.1. *Une fonction $y : I \rightarrow \mathbb{R}$ est une solution du problème de Cauchy (2.2) de donnée initiale (t_0, y_0) ssi*

1- *y est continue et $\forall t \in I \implies (t, y(t)) \in I \times U$.*

2- *$\forall t \in I \implies y(t) = y_0 + \int_0^t f(\tau, y(\tau))d\tau$.*

Chapitre 3

Résolution numérique du problème de Cauchy

Dans de nombreux cas, il n'est pas possible d'obtenir une expression analytique de la solution de (PC), pour cela nous avons besoin d'utiliser les méthodes numériques pour approximer cette solution

3.1 Principe générale des méthodes numériques

Pour obtenir une approximation numérique de la solution $y(x)$ sur $I = [t_0, T]$, nous allons estimer la valeur de cette fonction en un nombre fini de points t_i $i = 0, \dots, n$ constituant les nœuds de maillage. La solution numérique discrète obtenue aux points t_i est notée y_i . La distance entre les abscisse notée h est appelé pas de discrétisation, ce pas est constant dans les méthodes simple mais il peut variable $h_i = t_{i+1} - t_i$.

Les techniques de résolution des EDO sont basés sur

- L'approximation géométrique.
- Les formules d'intégration numérique (Trapèze, Simpson,...).
- Le développement de Taylor au voisinage de t_i .

3.1.1 Propriétés des méthodes numériques

Les propriété principale des méthodes numériques sont la consistance, la stabilité et la convergence ces propriétés permettant de relier la solution des équations continues à la solution numérique obtenue.

Remarque 3.1.1. *On peut ajouter à ces propriétés la notion de précision ainsi que des aspects informatiques comme la facilité de mise en œuvre, les coûts et mémoire.*

Consistance d'une méthode

La consistance est la propriété qui assure que la solution exacte de l'équation discrétisée tende vers la solution exacte de l'équation continue lorsque le pas de discrétisation tend vers 0.

Stabilité d'une méthode

C'est la propriété qui assure que la différence entre la solution numérique obtenue et la solution exacte des équations discrétisées reste bornée. La stabilité indique que si l'erreur augmente ou non au cours du calculs.

Ordre et précision d'une méthode

L'erreur de troncature ε_n est définie comme la différence entre la solution exacte $y(t_n)$ et l'approximation numérique y_n

$$\varepsilon_n = |y(t_n) - y_n| = O(h^p)$$

L'ordre de précision de la méthode est donnée par p .

Convergence et taux de convergence d'une méthode

Une méthode est convergente si lorsque le pas de discrétisation tend vers 0 la solution numérique tend vers la solution exacte de l'équation continue. Une méthode est convergente à l'ordre p si

$$\lim_{n \rightarrow \infty} \max_i |\varepsilon_i| = O(h^p)$$

Lemme 3.1.1. *Une méthode stable et consistante est convergente.*

3.1.2 Les principales méthodes numériques

Il ya deux classes des méthodes

Les méthodes à un pas

Pour ces méthodes le calcul de la valeur de y_{i+1} fait intervenir la valeur y_i obtenue à l'abscisse précédente. Les principales méthodes sont :

- Méthode d'Euler explicite et implicite.
- Méthode d'Euler modifiée.
- Méthode de Cauchy.
- Méthode de Crank Niicholson.
- Méthodes de Runge Kutta.

Les méthodes à pas multiples

Dans ces méthodes le calcul de y_{i+1} fait intervenir y_i, y_{i-1}, \dots . Les principales méthodes sont :

- Méthodes de Nystrom.
- Méthodes d'Adames- Bashforth.

-Méthodes de Gear.

Méthodes à un pas

la forme générale des méthodes explicites est :

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1} = y_i + h\Phi(t_i, y_i, h) \end{cases}$$

où Φ définit la méthode utilisée. Tandis que les méthodes à un pas implicites sont données par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1} = y_i + h\Phi(t_i, y_i, y_{i+1}, h) \end{cases}$$

Remarque 3.1.2. Ces méthodes sont obtenues en intégrant l'équation différentielle et en utilisant des formules d'intégration numérique pour le second membre.

Méthodes d'Euler explicite et implicite

Euler explicite

L'algorithme d'Euler explicite est donné par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1} = y_i + hf(t_i, y_i) \end{cases}$$

L'interprétation de la méthode

On peut l'interpréter de plusieurs manières

- 1 Géométriquement : La méthode revient à remplacer en chaque point t_i la courbe solution par sa tangente au point t_i .
- 2 Via le développement de Taylor : La méthode résulte du développement de Taylor d'ordre 1 de la fonction y au point t_i .

Euler implicite

L'algorithme est donné par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1} = y_i + hf(t_{i+1}, y_{i+1}) \end{cases}$$

Exemple 3.1.1. Soit le problème de Cauchy suivant

$$\begin{cases} y'(t) = y(t) - \frac{2t}{y(t)} & t \in [0, 1] \\ y(0) = 1 \end{cases}$$

On discrétise le domaine $[0, 1]$, on pose $N = 10$, $h = 0.1$, $t_0 = 0$, $t_i = ih$ et $t_N = 1$, y_i sont les solutions approchées aux points t_i . La fonction

$$f(t_i, y_i) = y_i - \frac{2t_i}{y_i}$$

l'algorithme d'Euler explicite est

$$\begin{cases} y_0 = 1 \\ y_{i+1} = y_i + hf(t_i, y_i) = y_i + h(y_i - \frac{2t_i}{y_i}) \end{cases}$$

on a $y_1 = y_0 + 0.1(y_0 - \frac{2t_0}{y_0}) = 1 + 0.1(1 + \frac{2*0}{1}) = 1.1$. L'algorithme implicite donne l'équation $0.9y_1^2 - y_1 + 0.02 = 0$, dans ce cas $y_1 = 1.098$.

La solution exacte de ce problème est $y(t) = \sqrt{2t+1}$, la valeur exacte de y au point $t_1 = 0.1$ est $y(0.1) = \sqrt{1.2} = 1.095$. L'erreur $E_{exp} = 0.005$ et $E_{imp} = 0.0025$.

Méthode d'Euler amélioré

L'algorithme explicite de cette méthode est donné par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1}^* = y_i + \frac{h}{2} f(t_i, y_i) \\ y_{i+1} = y_i + h f(t_i + \frac{h}{2}, y_{i+1}^*) \end{cases}$$

Géométriquement la méthode consiste à remplacer dans la méthode d'Euler la pente de la tangente en (t_i, y_i) par la valeur corrigée au milieu de l'intervalle $[t_i, t_{i+1}]$.

Exemple 3.1.2. D'après l'exemple précédent on trouve les résultats suivants :

$$\begin{cases} y_0 = 1 \\ y_{i+1}^* = y_i + \frac{h}{2} f(t_i, y_i) = y_i + \frac{h}{2} (y_i - \frac{2t_i}{y_i}) \\ y_{i+1} = y_i + h f(t_i + \frac{h}{2}, y_{i+1}^*) = y_i + h (y_{i+1}^* - 2(\frac{t_i + h/2}{y_{i+1}^*})) \end{cases}$$

donc $y_1^* = y_0 + \frac{0.1}{2}(y_0 - \frac{2t_0}{y_0}) = 1 + 0.05 = 1.05 \Rightarrow y_1 = 1 + 0.1(1.05 - \frac{0.1}{1.05}) = 1.0955$. L'erreur est donné par : $E = 0.000034$.

Méthode d'Euler Cauchy (RK2)

L'algorithme explicite est donné par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1}^* = y_i + h f(t_i, y_i) \\ y_{i+1} = y_i + \frac{h}{2} [f(t_i, y_i) + f(t_{i+1}, y_{i+1}^*)] \end{cases}$$

donc $y_1^* = 1.1$ et $y_1 = 1 + 0.05(1 + (1.1 - 2\frac{0.1}{1.1})) = 1.096$. L'erreur est donné par $E = 0.00046$.

Méthode de Crank-Nicholson

L'algorithme est donné par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1} = y_i + \frac{h}{2} [f(t_i, y_i) + f(t_{i+1}, y_{i+1})] \end{cases}$$

Il est obtenu en utilisant la formule d'intégration de Trapèze .

Remarque 3.1.3. La méthode de Crank-Nicholson c'est une méthode implicite d'ordre 2.

Exemple 3.1.3. De l'exemple précédent, on trouve

$$\begin{cases} y_0 = 1 \\ y_{i+1} = y_i + \frac{h}{2} [(y_i - \frac{2t_i}{y_i}) + (y_{i+1} - \frac{2t_{i+1}}{y_{i+1}})] \end{cases}$$

donc $y_1 = 1 + 0.05(1 + (y_1 - \frac{0.2}{y_1}))$, pour déterminer y_1 il faut résoudre une équation algébrique d'ordre 2 par rapport à y_1 . Après les calculs, on trouve $y_1 = 1.05$.

Méthodes de Runge Kutta explicite et implicite

Runge kutta explicite d'ordre 4 RK4

C'est une méthode explicite d'ordre 4, elle est obtenue par l'utilisation de la formule d'intégration de Simpson, elle s'écrit comme suit :

$$\left\{ \begin{array}{l} y_0 \text{ donné} \\ K_1 = f(t_i, y_i) \\ K_2 = f(t_i + \frac{h}{2}, y_i + \frac{K_1}{2}) \\ K_3 = f(t_i + \frac{h}{2}, y_i + \frac{K_2}{2}) \\ K_4 = f(t_i + h, y_i + K_3) \\ y_{i+1} = y_i + \frac{h}{6}[K_1 + 2(K_2 + K_3) + K_4] \end{array} \right.$$

Dans l'exemple précédent $y_0 = 1$, on calcul les constantes K_1, K_2, K_3 et K_4 on trouve :
 $K_1 = 0.1, K_2 = 0.095, K_3 = 0.095$ et $K_4 = 0.091$ donc $y_1 = 1.095 = y_{ex}(0.1)$.

Runge Kutta implicite

La méthode de Runge Kutta implicite est d'ordre 3, elle s'écrit comme suit :

$$\left\{ \begin{array}{l} y_0 \text{ donné} \\ K_1 = hf(t_i + \frac{h}{3}, y_i + \frac{h}{12}(5K_1 + K_2)) \\ K_2 = hf(t_i + h, y_i + \frac{h}{4}(3K_1 + K_2)) \\ y_{i+1} = y_i + \frac{1}{4}[3K_1 + K_2] \end{array} \right.$$

Dans ce cas pour déterminer y_{i+1} , il faut résoudre un système par rapport à K_1 et K_2

3.2 Analyse des méthodes à un pas (convergence, stabilité et consistance)

la forme générale des méthodes explicites à un pas est donnée par

$$\left\{ \begin{array}{l} y_0 \text{ donné} \\ y_{i+1} = y_i + h\Phi(t_i, y_i, h) \end{array} \right. \tag{3.1}$$

Définition 3.2.1. Le schéma (3.1) du problème (PC) converge si $\forall y_0 \in \mathbb{R}$

$$\lim_{h \rightarrow 0} \max_{0 \leq k \leq N} |y_k - y(t_k)| = 0$$

avec $y(t_k)$ est la solution du problème (PC).

Définition 3.2.2. (Stabilité (zéro-stabilité)) Le schéma (3.1) est stable s'il existe des constantes M_1 et M_2 indépendantes de h telles que $\forall y_{0h}, z_{0h} \in \mathbb{R}$ les suites y_k, z_k définies par

$$\begin{aligned} y_{k+1} &= y_k + h\Phi(t_k, y_k, h) & y_0 &= y_{0h} \\ z_{k+1} &= z_k + h\Phi(t_k, z_k, h) + \varepsilon_k & z_0 &= z_{0h} \end{aligned}$$

verifient

$$\max_{0 \leq k \leq N} |y_k - z_k| \leq M_1 |y_{0h} - z_{0h}| + M_2 \sum_{k=0}^{N-1} |\varepsilon_k|$$

Définition 3.2.3. (Consistance d'un schéma) On dit que le schéma (3.1) est consistant avec le problème (PC) si $\forall y$ solution du problème (PC) on a

$$\max_{0 \leq k \leq N} \left| \frac{y(t_{k+1}) - y(t_k)}{h} - \Phi(t_k, y(t_k), h) \right| \xrightarrow{h \rightarrow 0} 0$$

Remarque 3.2.1. La quantité $\tau_k = \frac{y(t_{k+1}) - y(t_k)}{h} - \Phi(t_k, y(t_k), h)$ s'appelle l'erreur de troncature locale, $\max_{0 \leq k \leq N} |\tau_k|$ est l'erreur de troncature globale.

Théorème 3.2.1. Si le schéma est stable et consistant, alors il est converge.

Critères de consistance et de stabilité

Théorème 3.2.2. Soient l'équation différentielle $y' = f(t, y)$ avec f continue de $[t_0, T] \times \mathbb{R}$ et lipschitzienne par rapport à y . Le schéma $y_{k+1} = y_k + h\Phi(t_k, y_k, h)$ avec Φ continue de $[t_0, T] \times \mathbb{R} \times [0, h_0]$ dans \mathbb{R} est consistant ssi

$$\Phi(t, y, 0) = f(t, y) \quad \forall (t, y) \in [t_0, T] \times \mathbb{R}$$

Théorème 3.2.3. Soit le schéma à un pas $y_{k+1} = y_k + h\Phi(t_k, y_k, h)$ une condition suffisante de stabilité est donnée par : Il existe une constante k indépendante de h , telle que

$$|\Phi(t, y, h) - \Phi(t, z, h)| \leq k|y - z| \quad \forall (t, y, z, h) \in [t_0, T] \times \mathbb{R}^2 \times [0, h_0]$$

Ordre d'une méthode à un pas

Définition 3.2.4. Le schéma à un pas $y_{k+1} = y_k + h\Phi(t_k, y_k, h)$ de discrétisation de l'équation $y' = f(t, y)$ est d'ordre p si $\max_{0 \leq k \leq N} |\tau_k| \leq Ch^p$ tel que la constante C ne dépendant que de y et de Φ .

Critère pour l'ordre p

Pour f suffisamment différentiable, posons

$$\begin{cases} f^{(0)} = f \\ f^{(1)} = \frac{\partial f^{(0)}}{\partial t} + \frac{\partial f^{(0)}}{\partial y} f \\ \vdots \\ f^{(k)} = \frac{\partial f^{(k-1)}}{\partial t} + \frac{\partial f^{(k-1)}}{\partial y} f \end{cases}$$

Exercice 3.2.1. Vérifier que si y est une solution de $y' = f(t, y)$, on a

$$f^{(k)}(t, y(t)) = y^{(k+1)}(t) = \frac{\partial^k}{\partial t^k} f(t, y(t))$$

Théorème 3.2.4. Si $f \in C^p([t_0, T] \times \mathbb{R})$ et $\Phi, \frac{\partial \Phi}{\partial h}, \dots, \frac{\partial^p \Phi}{\partial h^p}$ sont continues sur $[t_0, T] \times \mathbb{R} \times [0, h_0]$, alors la méthode à un pas

$$y_{k+1} = y_k + h\Phi(t_k, y_k, h)$$

est d'ordre p ssi $\forall (t, y) \in [t_0, T] \times \mathbb{R}$

$$\begin{cases} \Phi(t, y, 0) = f(t, y) \\ \frac{\partial \Phi}{\partial h}(t, y, 0) = \frac{1}{2} f^{(1)}(t, y) \\ \vdots \\ \frac{\partial^{p-1} \Phi}{\partial h^{p-1}}(t, y, 0) = \frac{1}{p} f^{(p-1)}(t, y) \end{cases}$$

et la constante C dans la définition de l'ordre est donnée par

$$C = \frac{1}{(p+1)!} \max_t |f^{(p)}(t, y(t))| + \frac{1}{p!} \max_{t,h} \left| \frac{\partial^p \Phi}{\partial h^p}(t, y(t), h) \right|$$

3.2.1 Applications sur les méthodes numériques précédentes

Euler explicite le schéma d'Euler explicite est

$$\begin{cases} y_0 & \text{donné} \\ y_{i+1} = y_i + hf(t_i, y_i) \end{cases}$$

on a $\Phi(t, y, h) = f(t, y) \Rightarrow \Phi(t, y, 0) = f(t, y)$ donc le schéma est consistant au moins d'ordre 1, de plus il est stable car $\Phi = f$ et f est lipschitzienne par rapport à y , d'autre part Φ ne dépend pas de h , donc $\frac{\partial \Phi}{\partial h}(t, y, 0) \neq \frac{1}{2} f^{(1)}(t, y)$. D'après le théorème de convergence la méthode converge avec un ordre égal à 1.

La méthode d'Euler implicite

Le schéma d'Euler implicite est

$$\begin{cases} y_0 & \text{donné} \\ y_{i+1} = y_i + hf(t_{i+1}, y_{i+1}) = y_i + h\Phi(t_{i+1}, y_{i+1}, h) \end{cases}$$

donc

$$\Phi(t, y, h) = f(t, y) \Rightarrow \Phi(t, y, 0) = f(t, y)$$

Donc le schéma est consistant au moins d'ordre 1. comme le schéma est stable car il est implicite donc le schéma converge. On peut montrer la consistence en utilisant la définition. En effet, le fait $\Phi(t_{i+1}, y_{i+1}, h) = f(t_{i+1}, y_{i+1}) = y'(t_{i+1})$, donc le développement de Taylor sera au point t_{i+1}

$$y(t_i) = y(t_{i+1} - h) = y(t_{i+1}) - hy'(t_{i+1}) + \frac{h^2}{2} y^{(2)}(\tau) \quad \text{avec } \tau \in]t_i, t_{i+1}[$$

ce qui implique

$$\begin{aligned} \frac{y(t_{i+1}) - y(t_i)}{h} - y'(t_{i+1}) &= -\frac{h}{2} y^{(2)}(\tau) \\ \Leftrightarrow \max_i \left| \frac{y(t_{i+1}) - y(t_i)}{h} - y'(t_{i+1}) \right| &\leq Mh \end{aligned}$$

Donc le schéma est consistant d'ordre 1 d'après le théorème de convergence le schéma converge d'ordre 1.

La méthode d'Euler modifiée

Dans cette méthode $\Phi(t, y, h) = f(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)) \Rightarrow \Phi(t, y, 0) = f(t, y)$, donc la méthode est consistante au moins d'ordre 1. Pour déterminer l'ordre en effet

$$\frac{\partial \Phi}{\partial h}(t, y, h) = \frac{\partial}{\partial h} f\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) = \frac{1}{2} \left(\frac{\partial f^{(0)}}{\partial t} \left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) + \frac{\partial f^{(0)}}{\partial y} \left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) f(t, y) \right)$$

donc

$$\frac{\partial \Phi}{\partial h}(t, y, 0) = \frac{1}{2} f^{(1)}(t, y)$$

D'après un calcul simple, on trouve

$$\frac{\partial^2 \Phi}{\partial h^2}(t, y, 0) = \frac{1}{4}f^{(2)}(t, y) \neq \frac{1}{3}f^{(2)}(t, y)$$

ce qui implique que la méthode est exactement d'ordre 2.

D'autre part on a

$$\begin{aligned} |\Phi(t, y, h) - \Phi(t, z, h)| &= |f(t + \frac{h}{2}, y + \frac{h}{2}f(t, y)) - f(t + \frac{h}{2}, z + \frac{h}{2}f(t, z))| \\ &\leq L|y + \frac{h}{2}f(t, y) - (z + \frac{h}{2}f(t, z))| \\ &\leq L|y - z| + L\frac{h}{2}|f(t, y) - f(t, z)| \\ &\leq L(1 + L\frac{h_0}{2})|y - z| \quad \text{avec } h \in [0, h_0] \\ &\leq C|y - z| \end{aligned}$$

Combinant la consistance et la stabilité, on trouve que la méthode converge d'ordre 2.

Méthode d'Euler Cauchy

On a

$$y_{i+1} = y_i + \frac{h}{2}[f(t_i, y_i) + f(t_{i+1}, y_i + hf(t_i, y_i))]$$

donc

$$\Phi(t, y, h) = \frac{1}{2}[f(t, y) + f(t+h, y+hf(t, y))] \Rightarrow \Phi(t, y, 0) = f(t, y)$$

ce qui donne que la méthode est consistante au moins d'ordre 1, de plus

$$\frac{\partial \Phi}{\partial h}(t, y, h) = \frac{1}{2} \frac{\partial}{\partial h} [f(t, y) + f(t+h, y+hf(t, y))] = \frac{1}{2} \left(\frac{\partial f^{(0)}}{\partial t}(t+h, y+hf(t, y)) + \frac{\partial f^{(0)}}{\partial y}(t+h, y+hf(t, y))f(t, y) \right)$$

ce qui implique que

$$\frac{\partial \Phi}{\partial h}(t, y, 0) = \frac{1}{2}f^{(1)}(t, y) \quad \text{et} \quad \frac{\partial^2 \Phi}{\partial h^2}(t, y, 0) \neq \frac{1}{3}f^{(2)}(t, y)$$

On conclut que la méthode est exactement d'ordre 2. On peut vérifier facilement que Φ est Lipschitzienne par rapport à y , d'après le théorème de convergence la méthode converge avec un ordre égale à 2.

Étude de la convergence de la méthode de Crank-Nicholson

Le schéma de Crank-Nicholson est donné par

$$\begin{cases} y_0 \text{ donné} \\ y_{i+1} = y_i + \frac{h}{2}[f(t_i, y_i) + f(t_{i+1}, y_{i+1})] = y_i + h\Phi(t_i, y_i, y_{i+1}, h) \end{cases}$$

— La consistance et son ordre

Dans cette méthode on utilise la définition de consistance, donc il faut calculer la quantité

$$\left| \frac{y(t_{i+1}) - y(t_i)}{h} - \Phi(t_i, y(t_i), y(t_{i+1}), h) \right|$$

avec la fonction $\Phi(t_i, y(t_i), y(t_{i+1}), h) = \frac{1}{2}[f(t_i, y(t_i)) + f(t_{i+1}, y(t_{i+1}))] = \frac{1}{2}[y'(t_i) + y'(t_{i+1})]$. Pour cela on développe la fonction $y'(t_i)$ au point t_{i+1} , on aura

$$y'(t_i) = y'(t_{i+1} - h) = y'(t_{i+1}) - hy''(t_{i+1}) + \frac{h^2}{2}y^{(3)}(t_{i+1}) - \frac{h^3}{3!}y^{(4)}(\tau_1) \quad \text{avec } \tau_1 \in]t_i, t_{i+1}[$$

donc

$$\begin{aligned} \frac{1}{2}[y'(t_i) + y'(t_{i+1})] &= \frac{1}{2}[y'(t_{i+1}) - hy''(t_{i+1}) + \frac{h^2}{2}y^{(3)}(t_{i+1}) - \frac{h^3}{2}y^{(4)}(\tau_1) + y'(t_{i+1})] \\ &= y'(t_{i+1}) + \frac{1}{2}[-hy''(t_{i+1}) + \frac{h^2}{2}y^{(3)}(t_{i+1}) - \frac{h^3}{3!}y^{(4)}(\tau_1)] \end{aligned} \quad (3.2)$$

D'autre part on peut développer la fonction $y(t_i)$ au point t_{i+1} , on obtient

$$y(t_i) = y(t_{i+1} - h) = y(t_{i+1}) - hy'(t_{i+1}) + \frac{h^2}{2}y^{(2)}(t_{i+1}) - \frac{h^3}{3!}y^{(3)}(t_{i+1}) + \frac{h^4}{4!}y^{(4)}(\tau_2) \quad \text{avec } \tau_2 \in]t_i, t_{i+1}[$$

donc

$$\frac{y(t_{i+1}) - y(t_i)}{h} = y'(t_{i+1}) - \frac{h}{2}y^{(2)}(t_{i+1}) + \frac{h^2}{3!}y^{(3)}(t_{i+1}) - \frac{h^3}{4!}y^{(4)}(\tau_2) \quad (3.3)$$

de (3.2) et (3.3) on trouve

$$\begin{aligned} \frac{y(t_{i+1}) - y(t_i)}{h} - \Phi(t_i, y(t_i), y(t_{i+1}), h) &= y'(t_{i+1}) - \frac{h}{2}y^{(2)}(t_{i+1}) + \frac{h^2}{3!}y^{(3)}(t_{i+1}) - \frac{h^3}{4!}y^{(4)}(\tau_2) - [y'(t_{i+1}) \\ &\quad + \frac{1}{2}(-hy''(t_{i+1}) + \frac{h^2}{2}y^{(3)}(t_{i+1}) - \frac{h^3}{3!}y^{(4)}(\tau_1))] \\ &= \frac{h^2}{6}y^{(3)}(t_{i+1}) - \frac{h^2}{4}y^{(3)}(t_{i+1}) + \frac{h^3}{12}y^{(4)}(\tau_1) - \frac{h^3}{24}y^{(4)}(\tau_2) \\ &= h^2(\frac{1}{6}y^{(3)}(t_{i+1}) - \frac{1}{4}y^{(3)}(t_{i+1})) + h^3(\frac{1}{12}y^{(4)}(\tau_1) - \frac{1}{24}y^{(4)}(\tau_2)) \end{aligned}$$

Donc

$$\max_i \left| \frac{y(t_{i+1}) - y(t_i)}{h} - \Phi(t_i, y(t_i), y(t_{i+1}), h) \right| \leq Ch^2$$

avec $C = \frac{5}{12}\|y^{(3)}\|_\infty$, ce qui implique que la méthode est consistante d'ordre 2, comme elle est stable, on trouve qu'elle est convergente.

3.3 Méthodes à pas multiples

Méthodes d'Adams Bashforth

Sont des méthodes basées sur les techniques d'intégration numérique, dont l'expression générale d'elles à q pas est :

$$y_{i+1} = y_i + h \sum_{j=0}^{q-1} \beta_{q,j}^i f(t_{i-j}, y_{i-j})$$

avec

$$\beta_{q,j}^i = \frac{1}{h} \int_{t_i}^{t_{i+1}} \mathcal{L}_{q,j}^i(t) dt$$

et $\mathcal{L}_{q,j}^i(t)$ sont les polynômes de base de l'interpolation de Lagrange de degré $q-1$ aux points $(t_{i-j}, f_{i-j})_{0 \leq j \leq q-1}$ et $f_{i-j} = f(t_{i-j}, y_{i-j})$, donc $\mathcal{L}_{q,j}^i(t) = \prod_{k=0, k \neq j}^{q-1} \frac{(t-t_{i-k})}{(t_{i-j}-t_{i-k})}$

Méthode d'Adams Bashforth à 2 pas

Pour $q = 2$, les points d'interpolations sont $(t_{i-j}, f_{i-j})_{0 \leq j \leq 1}$ c-à-d (t_i, f_i) et (t_{i-1}, f_{i-1}) , donc le schéma numérique est donné par :

$$\begin{cases} y_0 & \text{donné} \\ y_1 & \text{est calculé par une méthode à un pas} \\ y_{i+1} & = y_i + h[\beta_{2,0}^i f_i + \beta_{2,1}^i f_{i-1}] \end{cases}$$

avec

$$\beta_{2,0}^i = \frac{1}{h} \int_{t_i}^{t_{i+1}} \mathcal{L}_{2,0}^i(t) dt \quad \text{et} \quad \beta_{2,1}^i = \frac{1}{h} \int_{t_i}^{t_{i+1}} \mathcal{L}_{2,1}^i(t) dt$$

où

$$\mathcal{L}_{2,0}^i(t) = \frac{t-t_{i-1}}{t_i-t_{i-1}} = \frac{t-t_{i-1}}{h} \quad \text{et} \quad \mathcal{L}_{2,1}^i(t) = \frac{t-t_i}{t_{i-1}-t_i} = -\frac{t-t_i}{h}$$

donc

$$\beta_{2,0}^i = \frac{3}{2} \quad \text{et} \quad \beta_{2,1}^i = -\frac{1}{2}$$

La méthode d'Adams-Bashforth à 2 pas explicite est donnée par

$$\begin{cases} y_0 & \text{donné} \\ y_1 & \text{est calculé par une méthode à un pas} \\ y_{i+1} & = y_i + \frac{h}{2}[3f(t_i, y_i) - f(t_{i-1}, y_{i-1})] \end{cases}$$

Exercice 3.3.1. Montrer que pour $q = 3$ le schéma explicite d'Adams Bashforth à 3 pas est donnée par

$$\begin{cases} y_0 & \text{donné} \\ y_1, y_2 & \text{sont calculés par une méthode à un pas} \\ y_{i+1} & = y_i + \frac{h}{12}[23f(t_i, y_i) - 16f(t_{i-1}, y_{i-1}) + 5f(t_{i-2}, y_{i-2})] \end{cases}$$

Exemple 3.3.1. Si on prend le problème précédent

$$\begin{cases} y'(t) = y(t) - \frac{2}{y(t)} & t \in [0, 1] \\ y(0) = 1 \end{cases}$$

donc le schéma est

$$\begin{cases} y_0 = 1 \\ y_1 = y_1(\text{Euler explicite}) = 1.1 \\ y_2 = y_1 + \frac{0.1}{2}[3f(t_1, y_1) - f(t_0, y_0)] = 1.0877 \end{cases}$$

d'autre part la solution exacte au point $t = 0.2$ est $y(0.2) = \sqrt{1.4} = 1.183$

Les méthodes d'Adams -Moulton

Sont des méthodes implicites, dont la forme générale de ces méthodes à q pas est :

$$y_{i+1} = y_i + h \sum_{j=0}^q \beta_{q,j} f(t_{i+1-j}, y_{i+1-j})$$

avec

$$\beta_{q,j} = \frac{1}{h} \int_{t_i}^{t_{i+1}} \mathcal{L}_{q,j}(t) dt$$

et $\mathcal{L}_{q,j}(t)$ sont les polynômes de base de l'interpolation de Lagrange de degré q aux points $(t_{i+1-j}, f_{i+1-j})_{0 \leq j \leq q}$.

Par exemple la méthode d'Adams Moulton implicite à 1 pas est :

$$\begin{cases} y_0 & \text{donné} \\ y_{i+1} = y_i + \frac{h}{2} [f(t_i, y_i) + f(t_{i+1}, y_{i+1})] \end{cases}$$

et pour $q = 2$ la méthode est donnée par :

$$\begin{cases} y_0 & \text{donné} \\ y_1 & \text{est calculé par une méthode à un pas} \\ y_{i+1} = y_i + \frac{h}{12} [-f(t_{i-1}, y_{i-1}) + 8f(t_i, y_i) + 5f(t_{i+1}, y_{i+1})] \end{cases}$$

Remarque 3.3.1. Pour rendre les méthodes d'Adams-Moulton explicites, on remplace y_{i+1} "génant" par son estimation par la méthode d'Adams-Bashforth, cette technique construit des schémas explicites dits prédicteur-correcteur.

3.3.1 Analyse de convergence des méthodes à q pas

La forme générale des méthodes linéaires à q pas est

$$\begin{cases} \sum_{i=0}^q \alpha_i y_{k+i} = h \sum_{i=0}^q \beta_i f_{k+i} & 0 \leq k \leq N - q \\ y_i = y_{i,0} & 0 \leq i \leq q - 1 \end{cases} \quad (3.4)$$

avec $f_i = f(t_i, y_i)$ et $\alpha_q \neq 0$. On suppose que $|\alpha_0| + |\beta_0| \neq 0$ pour que la méthode soit à q pas.

Remarque 3.3.2. La méthode est explicite si $\beta_q = 0$, sinon elle est implicite.

Proposition 3.3.1. Si la méthode à q pas (3.4) est implicite (càd $\beta_q \neq 0$), elle admet une unique solution y_{k+q} dès que

$$h < \frac{|\alpha_q|}{L|\beta_q|}$$

avec L la constante de Lipschitz de f .

Convergence, stabilité, consistance et ordre

Définition 3.3.1. La méthode est convergente si $\forall y_0 \in \mathbb{R}$

$$\lim_{\substack{h \rightarrow 0 \\ 0 \leq k \leq N \\ y_{i,0} \rightarrow y_{0,0} \leq i \leq q-1}} \max |y_k - y(t_k)| = 0$$

$y(t)$ est la solution exacte.

Définition 3.3.2. (stabilité(zéro-stabilité)) La méthode (3.4) est stable s'il existe deux constantes M_1 et M_2 indépendantes de h telles que $\forall y_{i,0}, z_{i,0} \in \mathbb{R}$, tels que $0 \leq i \leq q-1$, et $\forall \varepsilon_k \in \mathbb{R}$ les suites $(y_k)_k, (z_k)_k$ définies par :

$$\left\{ \begin{array}{l} \sum_{i=0}^q \alpha_i y_{k+i} = h \sum_{i=0}^q \beta_i f_{k+i} \quad 0 \leq k \leq N-q \\ y_i = y_{i,0} \quad 0 \leq i \leq q-1 \\ \sum_{i=0}^q \alpha_i z_{k+i} = h \sum_{i=0}^q \beta_i \tilde{f}_{k+i} + \varepsilon_k \quad 0 \leq k \leq N-q \\ z_i = z_{i,0} \quad 0 \leq i \leq q-1 \end{array} \right.$$

avec $f_i = f(t_i, y_i)$ et $\tilde{f}_i = f(t_i, z_i)$ vérifient

$$\max_{0 \leq k \leq N} |y_k - z_k| \leq M_1 \max_{0 \leq i \leq q-1} |y_{i,0} - z_{i,0}| + M_2 \sum_{0 \leq k \leq N-q} |\varepsilon_k|$$

Définition 3.3.3. La méthode à q pas est consistante avec le problème de Cauchy (PC) si $\forall y(t)$ solution du problème (PC) on a

$$\max_{0 \leq k \leq N-q} \left| \frac{1}{h} \sum_{i=0}^q \alpha_i y(t_{k+i}) - \sum_{i=0}^q \beta_i f(t_{k+i}, y(t_{k+i})) \right| \xrightarrow{h \rightarrow 0} 0$$

Théorème 3.3.1. Soient le problème (PC) avec f continue, Lipschitzienne par rapport à y et la méthode à q pas (3.4) Si la méthode est stable et consistante elle est convergente.

Définition 3.3.4. (Méthode à q pas d'ordre p) La méthode à q pas est d'ordre p pour le problème (PC) si $\forall y(t)$ solution de classe C^{p+1} du problème (PC) on a

$$\max_{0 \leq k \leq N-q} \left| \frac{1}{h} \sum_{i=0}^q \alpha_i y(t_{k+i}) - \sum_{i=0}^q \beta_i f(t_{k+i}, y(t_{k+i})) \right| \leq Ch^p$$

avec C est une constante indépendante de h .

Critère de consistance et d'ordre p

Théorème 3.3.2. La méthode à q pas

$$\left\{ \begin{array}{l} \sum_{i=0}^q \alpha_i y_{k+i} = h \sum_{i=0}^q \beta_i f_{k+i} \quad 0 \leq k \leq N-q \\ y_i = y_{i,0} \quad 0 \leq i \leq q-1 \end{array} \right.$$

est d'ordre p ssi

$$\left\{ \begin{array}{l} \sum_{i=0}^q \alpha_i = 0 \quad \text{et} \quad \sum_{i=0}^q i \alpha_i = \sum_{i=0}^q \beta_i \quad \text{consistante d'ordre 1} \\ \sum_{i=0}^q i^2 \alpha_i = 2 \sum_{i=0}^q i \beta_i \\ \vdots \\ \sum_{i=0}^q i^p \alpha_i = p \sum_{i=0}^q i^{p-1} \beta_i \end{array} \right.$$

Proposition 3.3.2. (Constante d'erreur) Si la solution y du problème (PC) est de classe C^{p+2} et la méthode à q pas (3.4) est d'ordre p , l'erreur de troncature s'écrit

$$\varepsilon_{k+q} = \frac{1}{h} \sum_{i=0}^q \alpha_i y(t_{k+i}) - \sum_{i=0}^q \beta_i f(t_{k+i}, y(t_{k+i})) = C_{p+1} h^p y^{(p+1)}(t_k) + O(h^{p+1})$$

avec $C_{p+1} = \frac{1}{(p+1)!} \sum_{i=0}^q i^{p+1} \alpha_i - \frac{1}{p!} \sum_{i=0}^q i^p \beta_i$. On appelle constante d'erreur de la méthode la quantité

$$C = \frac{C_{p+1}}{\sum_{i=0}^q \beta_i} = \frac{C_{p+1}}{\rho'(1)}$$

avec $\rho(t) = \sum_{i=0}^q \alpha_i t^i$.

Critère de stabilité

Proposition 3.3.3. Soit la méthode à q pas

$$\begin{cases} \sum_{i=0}^q \alpha_i y_{k+i} = h \sum_{i=0}^q \beta_i f_{k+i} & 0 \leq k \leq N - q \\ y_i = y_{i0} & 0 \leq i \leq q - 1 \end{cases}$$

Une condition nécessaire de stabilité est que toutes les racines du polynôme ρ avec $\rho(t) = \sum_{i=0}^q \alpha_i t^i$ soient de module inférieur ou égal à 1, de plus les racines de module 1 étant simples.

Remarque 3.3.3. Si la méthode est stable on a $\rho'(1) \neq 0$.

Théorème 3.3.3. Si f est lipschitzienne par rapport à y la condition précédente du proposition (3.3.3) est une condition nécessaire et suffisante de stabilité pour la méthode à q pas (3.4).

Définition 3.3.5. (Stabilité forte et faible) On dit que la méthode est fortement stable si 1 est la seule racine de module 1 du polynôme ρ et faiblement stable si le polynôme ρ a plusieurs racines de module 1.

3.3.2 Application sur les méthodes précédentes

— La méthode d'Adams Bashforth à 2 pas

le schéma est donné par

$$\Leftrightarrow \begin{cases} y_0 \text{ donné} \\ y_1 \text{ est calculé par une méthode à un pas} \\ y_{i+1} = y_i + \frac{h}{2} [3f(t_i, y_i) - f(t_{i-1}, y_{i-1})] \\ y_0 \text{ donné } y_1 \text{ calculé} \\ y_{i+2} = y_{i+1} + \frac{h}{2} [3f_{i+1} - f_i] \\ \Leftrightarrow y_{i+2} - y_{i+1} = h [\frac{3}{2} f_{i+1} - \frac{1}{2} f_i] \\ \Leftrightarrow \alpha_2 y_{i+2} + \alpha_1 y_{i+1} + \alpha_0 y_i = h (\beta_2 f_{i+2} + \beta_1 f_{i+1} + \beta_0 f_i) \end{cases}$$

3.3. MÉTHODES À PAS MULTIPLES POUR LA SOLUTION NUMÉRIQUE DU PROBLÈME DE CAUCHY

— *La consistance*

Nous vérifions les conditions de consistance

$$\sum_{i=0}^2 \alpha_i = 1 - 1 = 0 \quad \text{et} \quad \sum_{i=0}^2 i \alpha_i = 1 = \sum_{i=0}^2 \beta_i$$

donc la méthode est consistante, pour déterminer l'ordre, on détermine p tel que

$$\sum_{i=0}^q i^p \alpha_i = p \sum_{i=0}^q i^{p-1} \beta_i$$

pour $p = 2$ on a

$$\sum_{i=0}^2 i^2 \alpha_i = 3 = 2 \sum_{i=0}^2 i \beta_i$$

pour $p = 3$ on a

$$\sum_{i=0}^2 i^3 \alpha_i = 7 \quad \text{et} \quad 3 \sum_{i=0}^2 i^2 \beta_i = \frac{9}{2} \neq 7$$

donc la méthode est consistante d'ordre 2

— *Stabilité*

La méthode est stable ssi toutes les racines de $\rho(t) = \sum_{i=0}^q \alpha_i t^i$ sont de module inférieur ou égale à 1 de plus les racines de module 1 sont simples.

$$\rho(t) = \sum_{i=0}^q \alpha_i t^i = t^2 - t = t(t-1)$$

les racines de $\rho(t)$ sont $t_1 = 0$ et $t_2 = 1$, donc la méthode est fortement stable.

— **La méthode d'Adams- Moulton à 2pas**

Le schéma est donné par :

$$\left\{ \begin{array}{l} y_0 \quad \text{donné} \\ y_1 \text{ est calculé par une méthode à un pas} \\ y_{i+1} = y_i + \frac{h}{12} [-f(t_{i-1}, y_{i-1}) + 8f(t_i, y_i) + 5f(t_{i+1}, y_{i+1}(t))] \\ \Leftrightarrow y_{i+2} - y_{i+1} + 0y_i = \frac{h}{12} [-f(t_{i-1}, y_{i-1}) + 8f(t_i, y_i) + 5f(t_{i+1}, y_{i+1})] \\ \Leftrightarrow \alpha_2 y_{i+2} + \alpha_1 y_{i+1} + \alpha_0 y_i = h(\beta_2 f_{i+2} + \beta_1 f_{i+1} + \beta_0 f_i) \end{array} \right.$$

— *La consistance*

On a

$$\sum_{i=0}^2 \alpha_i = 0 \quad \text{et} \quad \sum_{i=0}^2 i \alpha_i = 1 = \sum_{i=0}^2 \beta_i$$

donc la méthode est consistante

— *L'ordre de la méthode*

On a

$$\left\{ \begin{array}{l} \sum_{i=0}^2 i^2 \alpha_i = 3 = 2 \sum_{i=0}^2 i \beta_i \\ \sum_{i=0}^2 i^3 \alpha_i = 7 = 3 \sum_{i=0}^2 i^2 \beta_i \\ \sum_{i=0}^2 i^4 \alpha_i = 15, \quad \text{et} \quad 4 \sum_{i=0}^2 i^3 \beta_i = 16 \end{array} \right.$$

donc la méthode est d'ordre 3.

Chapitre 4

Méthode de différences finies

4.1 Introduction

Dans quelques cas très particuliers, il est impossible de calculer explicitement des solutions des différents modèles mathématiques. Il est donc nécessaire d'avoir recours au calcul numérique sur l'ordinateur pour estimer qualitativement et quantitativement ces solutions

Le principe de toutes les méthodes de résolution numérique des E.D.P.S est : d'obtenir des valeurs numériques discrètes (c-à-d en nombre fini) qui approchent (en un sens convenable à préciser) la solution exacte. Dans ce procédé il faut bien conscient de deux problèmes.

- 1- On ne calcule pas des solutions exactes mais approchées.
- 2- On discrétise le problème en représentant des fonctions par un nombre fini de valeurs c-à-d que l'on passe du continu au discret.

Il existe de nombreuses méthodes d'approximation numérique des solutions d'E.D.P. Nous présentons maintenant une des anciennes et plus simple, appelée méthode des différences finies.

4.1.1 Principe de la méthode

Le principe de la méthode des différences finies est de remplacer les dérivées par des différences finies en utilisant des formules de Taylor dans les quelles on néglige les restes.

4.2 Différence finies pour l'équation de la chaleur

Pour simplifier la présentation, nous nous limitons à la dimension 1 d'espace. Nous considérons l'équation de la chaleur dans le domaine borné $[0, 1]$.

$$\begin{cases} \partial_t u - \nu \partial_x^2 u = 0 & \text{pour } (x, t) \in (0, 1) \times \mathbb{R}^{+*} \\ u(0, x) = u_0(x) & \text{pour } x \in [0, 1] \\ u(t, 0) = u(t, 1) = 0 & \text{pour } t \in \mathbb{R}^{+*} \end{cases} \quad (4.1)$$

Pour discrétiser le domaine $[0, 1] \times \mathbb{R}^{+*}$, on introduit un pas d'espace $\Delta x = \frac{1}{N+1}$ avec $N \geq 1$ et un pas de temps $\Delta t > 0$, et on définit les nœuds d'un maillage régulière $(t_n, x_j) = (n\Delta t, j\Delta x) \quad \forall n \geq 0$ et $j \in \{0, \dots, N+1\}$, on note u_j^n la valeur d'une solution discrète approchée au point (t_n, x_j) et $u(t_n, x_j)$ la solution exacte de (4.1) au point (t_n, x_j) .

Pour discrétiser l'équation $\partial_t u - \nu \partial_x^2 u = 0$, en effet :

- On approche la dérivée seconde en espace, on effectue un développement de Taylor en x .

$$u(t, x - \Delta x) = u(t, x) - \Delta x u_x(t, x) + \frac{(\Delta x)^2}{2!} u_{xx}(t, x) - \frac{(\Delta x)^3}{3!} u_{xxx}(t, x) + \frac{(\Delta x)^4}{4!} u_{xxxx}(t, x) + \mathcal{O}((\Delta x)^6)$$

et

$$u(t, x + \Delta x) = u(t, x) + \Delta x u_x(t, x) + \frac{(\Delta x)^2}{2!} u_{xx}(t, x) + \frac{(\Delta x)^3}{3!} u_{xxx}(t, x) + \frac{(\Delta x)^4}{4!} u_{xxxx}(t, x) + \mathcal{O}((\Delta x)^6)$$

Combinant les deux équations précédentes, on obtient

$$u(t, x - \Delta x) + u(t, x + \Delta x) - 2u(t, x) = (\Delta x)^2 u_{xx}(t, x) + 2 \frac{(\Delta x)^4}{4!} u_{xxxx}(t, x) + \mathcal{O}((\Delta x)^6).$$

Si Δx est suffisamment petit, on trouve une bonne approximation du $u_{xx}(t, x)$ donnée par la formule suivante :

$$u_{xx}(t, x) \simeq \frac{u(t, x - \Delta x) + u(t, x + \Delta x) - 2u(t, x)}{(\Delta x)^2} \quad (4.2)$$

Remarque 4.2.1. On peut présenter la formule (4.2) au point (t_n, x_j) par

$$\begin{aligned} u_{xx}(t_n, x_j) &\simeq \frac{u(t_n, x_j - \Delta x) + u(t_n, x_j + \Delta x) - 2u(t_n, x_j)}{(\Delta x)^2} \\ &= \frac{u(t_n, x_{j-1}) + u(t_n, x_{j+1}) - 2u(t_n, x_j)}{(\Delta x)^2} \end{aligned}$$

donc le schéma approximatif du u_{xx} au point (t_n, x_j) est le suivant

$$u_{xx}(t_n, x_j) \simeq \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} \quad (4.3)$$

La formule (4.3) est dite schéma centré car elle est symétrique en j .

- Pour approcher la dérivée en temps, on fait la même chose. Examinons 3 formules naturelles

1- La différence finie centrée :

$$u_t(t_n, x_j) \simeq \frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t}$$

conduit au schéma complètement symétrique par rapport à n et j appelé schéma centré ou schéma de **Richardson**

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} - \nu \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} = 0. \quad (4.4)$$

2- La différence finie décentrée . On distingue deux cas

a- La différence finie décentrée amont :

$$u_t(t_n, x_j) \simeq \frac{u_j^n - u_j^{n-1}}{\Delta t}$$

qui conduit au schéma :

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} - \nu \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} = 0 \quad (4.5)$$

Ce schéma est appelé schéma d'Euler retrograde

b- La différence finie décentrée aval

$$u_t(t_n, x_j) \simeq \frac{u_j^{n+1} - u_j^n}{\Delta t}$$

qui conduit au schéma

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - \nu \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} = 0 \quad (4.6)$$

Remarques 4.2.1. 1- La différence principale entre ces deux derniers schémas est que (4.5) est dit implicite car il faut résoudre un système d'équations linéaires pour calculer les valeurs $(u_j^n)_{1 \leq j \leq N}$ en fonction des valeurs précédentes $(u_j^{n-1})_{1 \leq j \leq N}$, tandis que (4.6) est dit explicite, puisqu'il donne immédiatement les valeurs $(u_j^{n+1})_{1 \leq j \leq N}$ en fonction de $(u_j^n)_{1 \leq j \leq N}$.

2- On peut réécrire de manière équivalente (4.6) sous la forme :

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} - \nu \frac{u_{j+1}^{n-1} - 2u_j^{n-1} + u_{j-1}^{n-1}}{(\Delta x)^2} = 0$$

3- Dans les trois schémas que nous venons de définir, il y a bien sûr une donnée initiale pour démarrer les itérations en n . La donnée initiale est discrétisée par

$$u_j^0 = u_0(x_j) \text{ pour } j \in \{0, \dots, N+1\}.$$

4- Pour le schéma centré (4.4) il y a une difficulté au démarrage. Pour $n = 1$ on a aussi besoin de connaître les valeurs $(u_j^1)_{1 \leq j \leq N}$ qu'il faut donc calculer autrement par exemple, par l'un des deux autres schémas.

5- Les conditions aux limites du problème (4.1) peuvent être de plusieurs types, mais leur choix n'intervient pas dans la définition des schémas.

Dans ce problème, nous utilisons des conditions aux limites de **Dirichlet** $u(t, 0) = u(t, 1) = 0$ pour tout $t \in \mathbb{R}_+^*$ qui se traduisent en

$$u_0^n = u_{N+1}^n = 0 \text{ pour tout } n > 0$$

Par conséquent, à chaque pas des temps nous avons à calculer les valeurs $(u_j^n)_{1 \leq j \leq N}$ qui forment un vecteur de \mathbb{R}^N .

- 6- Il est facile de vérifier que le schéma implicite (4.5) est effectivement bien défini c-à-d qu'on peut calculer les valeurs u_j^n en fonction des u_j^{n-1} , en effet

$$\begin{aligned} \frac{u_j^n - u_j^{n-1}}{\Delta t} - v \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} &= 0 \\ \iff u_j^n - u_j^{n-1} - \frac{v\Delta t}{(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) &= 0 \\ \iff -\frac{v\Delta t}{(\Delta x)^2}u_{j-1}^n + (1 + 2\frac{v\Delta t}{(\Delta x)^2})u_j^n - \frac{v\Delta t}{(\Delta x)^2}u_{j+1}^n &= u_j^{n-1}. \end{aligned}$$

Si on pose $c = \frac{v\Delta t}{(\Delta x)^2}$, le dernier système (4.7) est équivalent à

$$\begin{cases} \text{Pour } j = 1 \iff -cu_0^n + (1 + 2c)u_1^n - cu_2^n = u_1^{n-1} \\ \text{Pour } j = 2 \iff -cu_1^n + (1 + 2c)u_2^n - cu_3^n = u_2^{n-1} \\ \vdots \\ \text{Pour } j = N \iff -cu_{N-1}^n + (1 + 2c)u_N^n - cu_{N+1}^n = u_N^{n-1} \end{cases}$$

matriciellement est donné par

$$\begin{pmatrix} (1 + 2c) & -c & 0 & \dots & 0 & 0 \\ -c & (1 + 2c) & -c & \dots & 0 & 0 \\ 0 & -c & (1 + 2c) & -c & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & -c \\ 0 & 0 & \dots & \dots & -c & (1 + 2c) \end{pmatrix} \begin{pmatrix} u_1^n \\ u_2^n \\ \vdots \\ \vdots \\ u_N^n \end{pmatrix} = \begin{pmatrix} u_1^{n-1} \\ u_2^{n-1} \\ \vdots \\ \vdots \\ u_N^{n-1} \end{pmatrix}$$

- 7- Si on fait une combinaison convexe de (4.5) et (4.6) pour $0 \leq \theta \leq 1$, on obtient le θ -schéma

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} + \theta \left(-v \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} \right) + (1 - \theta) \left(-v \frac{u_{j+1}^{n-1} - 2u_j^{n-1} + u_{j-1}^{n-1}}{(\Delta x)^2} \right) = 0 \quad (4.7)$$

- 8- Pour $\theta = 0$ dans (4.7) on trouve le schéma explicite (4.6), tandis que le schéma implicite (4.5) est obtenu si $\theta = 1$.
- 9- Le schéma (4.7) est implicite si $\theta \neq 0$.
- 10- Pour $\theta = \frac{1}{2}$ on obtient le schéma de **Crank-Nicholson**
- 11- Tous les schémas précédents sont dits à deux niveaux car ils font intervenir que deux indices de temps.
- 12- On peut trouver des schémas multi-niveaux, les plus populaires sont à trois niveaux, par exemple le schéma de **Richardson** (4.4).
- 13- Un des buts de l'analyse numérique va être de comparer et de sélectionner les meilleurs schémas suivant des critères de précision, de coût ou de robustesse.
- 14- S'il y a un second membre $f(t, x)$ dans l'équation de la chaleur (4.1), alors les schémas se modifient en remplaçant zéro au second membre par une approximation de $f(t, x)$ au point (t_n, x_j) . Par exemple si on choisit l'approximation $f(t_n, x_j)$ le schéma implicite (4.5) devient

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} - v \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} = f(t_n, x_j)$$

- 15- La collection des couples (n', j') qui interviennent dans l'équation discrète au point (t_n, x_j) est appelée le support du schéma. Plus le support est large plus le schéma est coûteux et difficile à programmer à cause des effets de bord c-à-d des cas où certains couples (n', j') sortent du domaine de calcul.
- 16- On peut remplacer les conditions aux limites de **Dirichlet** dans (4.1) par des conditions aux limites de **Neumann**, ou bien des conditions aux limites de **périodicité**. Si on prend les conditions de **Neumann**

$$u_x(t, 0) = u_x(t, 1) = 0$$

On peut décrire deux manières différentes de discrétisation du conditions

16-1 Une discrétisation du 1^{er} ordre

$$\frac{u_1^n - u_0^n}{\Delta x} = 0 \text{ et } \frac{u_{N+1}^n - u_N^n}{\Delta x} = 0$$

qui est permet d'éliminer les valeurs u_0^n et u_{N+1}^n et ne calculer que les N valeurs $(u_j^n)_{1 \leq j \leq N}$.

16-2 Une discrétisation du 2^{eme} ordre

$$\frac{u_1^n - u_{-1}^n}{2\Delta x} = 0 \text{ et } \frac{u_{N+2}^n - u_N^n}{2\Delta x} = 0$$

cette discrétisation nécessite l'ajoute 2 points x_{-1} et x_{N+2} , on élimine les valeurs u_{-1}^n et u_{N+2}^n correspondant à ces points, et il restent maintenant $N + 2$ valeurs à calculer $(u_j^n)_{0 \leq j \leq N+1}$.

4.3 Convergence du schéma numérique

L'étude de convergence d'un schéma numérique est basé sur l'étude de deux propriétés principales, à savoir la consistance et la stabilité. Commençons par la consistance.

4.3.1 Consistance

Les formules des schémas précédents résultent d'une approximation de l'équation par développement de Taylor, pour construire un lien entre cette approximation (l'équation de différences finies) et l'équation aux dérivées partielles, on introduit la notion de consistance et de précision.

De manière générale, un schéma aux différences finies est défini pour tous les indices possible n, j par la formule

$$F_{\Delta t, \Delta x}((u_{j+k}^{n+m})_{\substack{m^- \leq m \leq m^+ \\ k^- \leq k \leq k^+}}) = 0 \tag{4.8}$$

où les entiers m^+, m^-, k^+, k^- définissent la largeur du support du schéma.

Définition 4.3.1. Le schéma aux différences finies (4.8) est dit consistant avec l'équation aux dérivées partielles $F(u) = 0$, si pour toute solution $u(t, x)$ suffisamment régulière de cette équation l'erreur de troncature du schéma définie par :

$$F_{\Delta t, \Delta x}(u(t + m\Delta t, x + k\Delta x)_{\substack{m^- \leq m \leq m^+ \\ k^- \leq k \leq k^+}}) \tag{4.9}$$

tend vers 0 uniformément par rapport à (t, x) lorsque Δt et Δx tendent vers 0 indépendamment.

L'étude de consistance est toujours accompagnée de détermination de la précision, pour cela, il faut donner la définition suivante.

Définition 4.3.2. On dit que le schéma est précis à l'ordre p en espace et à l'ordre q en temps si l'erreur de troncature (4.9) tend vers 0 comme $\mathcal{O}((\Delta t)^q, (\Delta x)^p)$ lorsque Δt et Δx tendent vers 0.

Remarque 4.3.1. Pour calculer l'erreur de troncature d'un schéma en remplaçant u_{j+k}^{n+m} dans la formule (4.9) par $u(t + m\Delta t, x + k\Delta x)$.

Exemple 4.3.1. L'erreur de troncature du schéma explicite (4.6) est

$$\begin{aligned} & \frac{u(t + \Delta t, x) - u(t, x)}{\Delta t} - v \frac{u(t, x + \Delta x) - 2u(t, x) + u(t, x - \Delta x)}{(\Delta x)^2} \\ = & u_t(t, x) - v u_{xx}(t, x) + \frac{\Delta t}{2} u_{tt}(t, x) - v \frac{(\Delta x)^2}{12} u_{xxx}(t, x) + \mathcal{O}((\Delta t)^2, (\Delta x)^4) \end{aligned}$$

Le schéma explicite (4.6) est consistant, précis à l'ordre 1 en temps et 2 en espace.

4.4 Stabilité

L'étudier de la stabilité nécessite à définir une norme, pour cela nous reprenons les normes classiques sur \mathbb{R}^N que nous pondérons par le pas d'espace Δx . Soit $u^n = (u_j^n)_{1 \leq j \leq N}$ la solution approchée obtenue par le schéma numérique, on définit les normes suivantes :

$$\|u^n\|_p = \left(\sum_{j=1}^N \Delta x |u_j^n|^p \right)^{\frac{1}{p}} \quad \text{pour } 1 \leq p < \infty \quad (4.10)$$

Le cas $p = \infty$

$$\|u^n\|_\infty = \max_{1 \leq j \leq N} |u_j^n| \quad (4.11)$$

Définition 4.4.1. Un schéma de différence finies est dit stable pour les normes (4.10) et (4.11), s'il existe une constante $k > 0$ indépendante de Δx et Δt lorsque ces valeurs tendent vers zéro telle que

$$\|u^n\|_p \leq k \|u^0\|_p \quad \text{pour } 1 \leq p \leq \infty, \text{ et } n \geq 0 \quad (4.12)$$

quelle que soit la donnée initiale u^0 .

Remarque 4.4.1. Si (4.12) est satisfait pour certaines conditions sur Δt et Δx , on dit que le schéma est conditionnement stable.

Définition 4.4.2. Un schéma aux différences finies est dit linéaire si la formule $F_{\Delta t, \Delta x}(u_{j+k}^{n+m}) = 0$ qui le définit est linéaire par rapport à ces arguments u_{j+k}^{n+m} .

Remarque 4.4.2. La stabilité d'un schéma linéaire à deux niveaux est très facile à interpréter, car tout schéma linéaire à deux niveaux peut s'écrire sous la forme

$$u^{n+1} = Au^n$$

où A est une matrice, dite d'itération de \mathbb{R}^N .

Par exemple, pour le schéma explicite (4.6) la matrice A vaut

$$\begin{pmatrix} 1-2c & c & 0 & \cdots & 0 \\ c & 1-2c & c & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & c \\ 0 & \cdots & \cdots & c & 1-2c \end{pmatrix} \text{ avec } c = \frac{v\Delta t}{\Delta x}$$

tandis que pour le schéma implicite (4.5) la matrice A est l'inverse de la matrice

$$\begin{pmatrix} 1+2c & -c & 0 & \cdots & 0 \\ -c & 1+2c & -c & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -c \\ 0 & \cdots & \cdots & -c & 1+2c \end{pmatrix}$$

A l'aide de cette matrice d'itération, on a $u^n = A^n u^0$, et par conséquent la stabilité du schéma est équivalente à

$$\|A^n u^0\| \leq k \|u^0\| \quad \forall n \geq 0, \forall u^0 \in \mathbb{R}^N$$

qui veut dire que la suite des puissances de A est bornée.

4.4.1 Stabilité en norme L^∞

La stabilité en norme L^∞ est très liée avec le principe du maximum discret.

Définition 4.4.3. Un schéma aux différences finies vérifie le principe du maximum discret si pour tout $n \geq 0$ et tout $0 \leq j \leq N+1$, on a

$$\min\left(0, \min_{0 \leq j \leq N+1} u_j^0\right) \leq u_j^n \leq \max\left(0, \max_{0 \leq j \leq N+1} u_j^0\right)$$

quelle que soit la donnée initiale.

Remarque 4.4.3. Dans la définition précédente, les inégalités tiennent compte non seulement du minimum et du maximum de u^0 mais aussi de zéro qui est la valeur imposée au bord par les conditions aux limites de Dirichlet. Cela est nécessaire si la donnée initiale u^0 ne vérifie pas les conditions aux limites de Dirichlet (ce qui n'est pas exigé), et unitil dans le cas contraire.

Lemme 4.4.1. Le schéma explicite (4.6) est stable en norme L^∞ ssi la condition $(2v\Delta t / (\Delta x)^2) < 1$ est satisfaite. Le schéma implicite (4.5) est stable en norme L^∞ quelque soit les pas de temps Δt et d'espace Δx (on dit qu'il est inconditionnellement stable).

Preuve 1. Exercice

4.4.2 Stabilité en norme L^2

La norme L^2 se prête très bien à l'étude de la stabilité grâce à l'outil très puissant de l'analyse de Fourier. Pour ce faire, nous supposons désormais que les conditions aux limites pour l'équation de la chaleur de périodicité, qui s'écrivent $u(t, x + 1) = u(t, x)$ pour tout $x \in [0, 1]$ et pour tout $t \geq 0$. Pour les schémas numériques ils conduisent aux $u_0^n = u_{N+1}^n$ pour tout $n \geq 0$, et plus généralement $u_j^n = u_{N+1+j}^n$. Il reste donc à calculer $N + 1$ valeurs de u_j^n .

A chaque vecteur $u^n = (u_j^n)_{0 \leq j \leq N}$ on associe une fonction $u^n(x)$, constante par morceaux, périodique de période 1, définie sur $[0, 1]$ par :

$$u^n(x) = u_j^n \text{ si } x_{j-\frac{1}{2}} < x < x_{j+\frac{1}{2}}$$

avec $x_{j \pm \frac{1}{2}} = x_j \pm \frac{1}{2} \Delta x$ pour $0 \leq j \leq N$ et $x_{-\frac{1}{2}} = 0, x_{N+1+\frac{1}{2}} = 1$.

La fonction $u^n \in L^2(0, 1)$ (on connaît que $L^2(0, 1)$ est un espace de Hilbert et que $\{e^{2i\pi kx} / k \in \mathbb{Z}\}$ est une base hilbertienne de $L^2(0, 1)$). Donc la fonction u^n peut se décomposer en une somme de Fourier et on a

$$u^n(x) = \sum_{k \in \mathbb{Z}} \hat{u}^n(k) \exp(2i\pi kx)$$

avec

$$\hat{u}^n(k) = \int_0^1 u^n(x) \exp(-2i\pi kx) dx$$

et la formule de Plancherel

$$\int_0^1 |u^n(x)|^2 dx = \sum_{k \in \mathbb{Z}} |\hat{u}^n(k)|^2$$

Remarque 4.4.4. 1- Remarquons que même si u^n est une fonction réelle, les coefficients $\hat{u}^n(k)$ de la série de Fourier sont complexes.

2- Une propriété importante pour la suite de la transformée de Fourier des fonction périodiques est la suivante : si on note $v^n(x) = u^n(x + \Delta x)$ alors $\hat{v}^n(k) = \hat{u}^n(k) \exp(2i\pi k \Delta x)$

Exemple 4.4.1. Expliquons maintenant la méthode sur le schéma explicite (4.6).

On peut réécrire ce schéma, pour $x \in (0, 1)$ comme suit :

$$\frac{u^{n+1}(x) - u^n(x)}{\Delta t} - \nu \frac{u^n(x + \Delta x) - 2u^n(x) + u^n(x - \Delta x)}{(\Delta x)^2} = 0 \quad (4.13)$$

ce qui équivaut à

$$u^{n+1}(x) = u^n(x) + \nu \Delta t \frac{u^n(x + \Delta x) - 2u^n(x) + u^n(x - \Delta x)}{(\Delta x)^2}$$

Par l'application de la transformation de Fourier, il vient :

$$\begin{aligned} \hat{u}^{n+1}(k) &= \hat{u}^n(k) + \frac{\nu \Delta t}{(\Delta x)^2} [\hat{u}^n(k) \exp(2i\pi k \Delta x) - 2\hat{u}^n(k) + \hat{u}^n(k) \exp(-2i\pi k \Delta x)] \\ &= [1 + \frac{\nu \Delta t}{(\Delta x)^2} \exp(2i\pi k \Delta x) - 2 + \exp(-2i\pi k \Delta x)] \hat{u}^n(k) \\ &= [1 + \frac{\nu \Delta t}{(\Delta x)^2} (2 \cos(2\pi k \Delta x) - 2)] \hat{u}^n(k) \end{aligned}$$

donc on aura

$$\begin{aligned} \hat{u}^{n+1}(k) &= 1 + \frac{v\Delta t}{(\Delta x)^2} \left[2(\cos^2(\pi k \Delta x) - \sin^2(\pi k \Delta x)) - 2(\cos^2(\pi k \Delta x) + \sin^2(\pi k \Delta x)) \right] \hat{u}^n(k) \\ &= \left[1 - 4 \frac{v\Delta t}{(\Delta x)^2} \sin^2(\pi k \Delta x) \right] \hat{u}^n(k) \\ &= A(k) \hat{u}^n(k) \end{aligned}$$

avec

$$A(k) = 1 - 4 \frac{v\Delta t}{(\Delta x)^2} \sin^2(\pi k \Delta x)$$

et

$$\hat{u}^{n+1}(k) = (A(k))^{n+1} \hat{u}^0(k) \quad k \in \mathbb{Z}$$

Le coefficient de Fourier $\hat{u}^n(k)$ est bornée lorsque $n \rightarrow \infty$ ssi le facteur d'amplification vérifie $|A(k)| \leq 1$ c-à-d

$$4 \frac{v\Delta t}{(\Delta x)^2} \sin^2(\pi k \Delta x) < 2 \iff \frac{v\Delta t}{(\Delta x)^2} \sin^2(\pi k \Delta x) \leq \frac{1}{2} \quad (4.14)$$

Si la condition $2v\Delta t \leq (\Delta x)^2$ est satisfaite, alors l'inégalité (4.14) est vraie quelque soit le mode de Fourier $k \in \mathbb{Z}$. En utilisant la formule de Plancherel on obtient

$$\|u^n\|_2^2 = \int_0^1 |u^n(x)|^2 dx = \sum_{k \in \mathbb{Z}} |\hat{u}^n(k)|^2 \leq \sum_{k \in \mathbb{Z}} |\hat{u}^0(k)|^2 = \int_0^1 |u^0(x)|^2 dx = \|u^0\|_2^2.$$

Ce qui n'est rien d'autre que la stabilité L^2 du schéma explicite.

Lemme 4.4.2. le schéma implicite (4.5) est stable en norme L^2 .

4.5 Schémas multiniveaux

Examinons comment les résultats précédents se généralisent aux schémas multiniveaux. On peut envisager des schémas à 3 niveaux où u^{n+1} dépend de u^n et u^{n-1} comme schéma de **Richardson**.

Remarquons que la définition de stabilité d'un schéma est indépendante de son nombre de niveaux. En effet u^{n+1} dépend linéairement de u^n et u^{n-1} , donc si on pose

$$U^n = \begin{pmatrix} u^n \\ u^{n-1} \end{pmatrix}$$

alors il existe deux matrices A_1 et A_2 d'ordre N tel que

$$U^{n+1} = AU^n = \begin{pmatrix} A_1 & A_2 \\ I_N & 0 \end{pmatrix} \begin{pmatrix} u^n \\ u^{n-1} \end{pmatrix}$$

où la matrice d'itération A est donc de taille $2N$. Et on a

$$U^{n+1} = AU^n = A^2U^{n-1} = A^3U^{n-2} = \dots = A^nU^1$$

et la stabilité est équivalente à

$$\|A^n\| = \sup_{y \neq 0, y \in \mathbb{R}^{2N}} \frac{\|Ay\|}{\|y\|} \leq k \quad \forall n \geq 1$$

Exemple 4.5.1. Si on prend le schéma de Richardson

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} - \nu \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} = 0 \quad (4.15)$$

Le schéma (4.15) précédent est instable en norme L^2 car : pour $x \in [0, 1]$, (4.15) s'écrit :

$$\frac{u^{n+1}(x) - u^{n-1}(x)}{2\Delta t} - \nu \frac{u^n(x + \Delta x) - 2u^n(x) + u^n(x - \Delta x)}{(\Delta x)^2} = 0 \quad (4.16)$$

la dernière égalité (4.16) est équivalente à

$$u^{n+1}(x) - u^{n-1}(x) + \frac{2\nu\Delta t}{(\Delta x)^2} (-u^n(x + \Delta x) + 2u^n(x) - u^n(x - \Delta x)) = 0$$

par l'application de la transformée de Fourier

$$\hat{u}^{n+1}(k) - \hat{u}^{n-1}(k) + \frac{2\nu\Delta t}{(\Delta x)^2} \left(-\hat{u}^n(k) \exp(2i\pi k \Delta x) + 2\hat{u}^n(k) - \hat{u}^n(k) \exp(-2i\pi k \Delta x) \right) = 0$$

autrement dit

$$\begin{aligned} \Leftrightarrow & \hat{u}^{n+1}(k) - \hat{u}^{n-1}(k) + \frac{2\nu\Delta t}{(\Delta x)^2} \hat{u}^n(k) (-\exp(2i\pi k \Delta x) + 2 - \exp(-2i\pi k \Delta x)) = 0 \\ \Leftrightarrow & \hat{u}^{n+1}(k) = \hat{u}^{n-1}(k) + \frac{2\nu\Delta t}{(\Delta x)^2} \hat{u}^n(k) (2 \cos(2k\pi \Delta x) - 2) \\ = & \hat{u}^{n-1}(k) + \frac{2\nu\Delta t}{(\Delta x)^2} \hat{u}^n(k) (2(1 - 2 \sin^2 k\pi \Delta x) - 2) \\ = & (-4 \frac{2\nu\Delta t}{(\Delta x)^2} \sin^2 k\pi \Delta x) \hat{u}^n(k) + \hat{u}^{n-1}(k) \end{aligned}$$

Donc

$$\begin{aligned} \hat{U}^{n+1}(k) &= \begin{pmatrix} \hat{u}^{n+1}(k) \\ \hat{u}^n(k) \end{pmatrix} = \begin{pmatrix} -4 \frac{2\nu\Delta t}{(\Delta x)^2} \sin^2 k\pi \Delta x & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{u}^n(k) \\ \hat{u}^{n-1}(k) \end{pmatrix} \\ &= A(k) \hat{U}^n(k) = (A(k))^n \hat{U}^1(k). \end{aligned}$$

Donc le coefficient de Fourier est borné lorsque $n \rightarrow \infty$ ssi la matrice d'amplification vérifie :

$$\|(A(k))^n\|_2 = \sup_{y \neq 0, y \in \mathbb{R}^{2N}} \frac{\|(A(k))^n y\|_2}{\|y\|_2} \leq k \quad \forall n \geq 0$$

où $\|u\|_2$ est la norme euclidienne dans \mathbb{R}^2 . En utilisant la formule de Plancherel on déduit

$$\|u^n\|_2^2 = \sum_{k \in \mathbb{Z}} |\hat{u}^n(k)|^2 \leq k \sum_{k \in \mathbb{Z}} \left(|\hat{u}^0(k)|^2 + |\hat{u}^1(k)|^2 \right) = \|U^0\|_2^2 + \|U^1\|_2^2 \quad (4.17)$$

Comme la matrice d'amplification $A(k)$ est symétrique réelle, on a la propriété

$$\|A(k)\|_2 = \rho(A(k))$$

et

$$\|(A(k))^n\|_2 = \|A(k)\|_2^n$$

où $\rho(A(k))$ désigne le rayon spectral de la matrice $A(k)$, donc l'inégalité (4.17) est satisfaite ssi $\rho(A(k)) \leq 1$. Les valeurs propres de $A(k)$ sont les racines du polynôme du 2^{ème} degré

$$\lambda^2 + \left(\frac{8\nu\Delta t}{(\Delta x)^2} \sin^2 k\pi\Delta x \right) \lambda - 1 = 0$$

qui admet toujours deux racines réelles distinctes dont le produit vaut -1 , l'une des deux racines est plus grand que 1 en valeur absolue et donc $\rho(A(k)) > 1$. Par conséquent le schéma de Richardson est inconditionnellement instable en norme L^2 .

4.6 Erreurs de troncature et stabilité de divers schémas pour l'équation de la chaleur

Schéma	Erreur de troncature	Stabilité
Explicite	$\mathcal{O}(\Delta t + (\Delta x)^2)$	stable L^2 et L^∞ si $(2\nu\Delta t) \leq (\Delta x)^2$
Implicite	$\mathcal{O}(\Delta t + (\Delta x)^2)$	stable L^2 et L^∞
Crank-Nicholson avec $\theta = \frac{1}{2}$	$\mathcal{O}((\Delta t)^2 + (\Delta x)^2)$	stable L^2
θ - schéma avec $\theta \neq \frac{1}{2}$	$\mathcal{O}(\Delta t + (\Delta x)^2)$	stable L^2 si $(2(1-2\theta)\nu\Delta t) \leq (\Delta x)^2$
Schéma de Duffort Frankel $\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} - \nu \frac{u_{j+1}^n - (u_j^{n+1} + u_j^{n-1} + u_{j-1}^n)}{(\Delta x)^2} = 0$	$\mathcal{O}\left(\left(\frac{\Delta t}{\Delta x}\right)^2 + (\Delta x)^2\right)$	stable L^2
Schéma de Gear $\frac{3u_j^{n+1} - 4u_j^n + u_j^{n-1}}{2\Delta t} - \nu \frac{u_{j+1}^{n+1} - 2u_j^n + u_{j-1}^{n+1}}{(\Delta x)^2} = 0$	$\mathcal{O}((\Delta t)^2 + (\Delta x)^2)$	stable L^2