

# Analyse Numérique des EDO

**Yasmina Daikh**

Université de Jijel

Faculté des Sciences Exactes et Informatique

Département de Mathématiques

B.P. 98, Ouled Aissa

18000 Jijel, Algérie

yasmina-daikh@univ-jijel.dz.

# Table des matières

<b>1</b>	<b>Rappels sur le cours d'équations différentielles</b>	<b>4</b>
1.1	Généralités . . . . .	4
1.2	Enoncé du théorème de Cauchy-Lipschitz . . . . .	5
1.3	Sur la régularité des solutions . . . . .	5
<b>2</b>	<b>Méthodes numériques à un pas</b>	<b>6</b>
2.1	Quelques exemples de méthodes à un pas . . . . .	8
2.1.1	Méthodes d'Euler . . . . .	8
2.1.2	Méthode de Taylor d'ordre $p$ . . . . .	10
2.2	Etude des méthodes à un pas . . . . .	11
2.3	Influence de l'ordre de la méthode sur l'erreur globale . . . . .	16
2.4	Méthodes de type Runge-Kutta . . . . .	17
2.4.1	Description de la méthode . . . . .	17
2.4.2	Exemples . . . . .	18
2.4.3	Stabilité des méthodes de Runge-Kutta . . . . .	20
2.4.4	Ordre des méthodes de Runge-Kutta . . . . .	21
2.5	Application aux systèmes d'équations différentielles . . . . .	23
2.5.1	Algorithme (RK4) . . . . .	23
2.5.2	Exemple . . . . .	24
<b>3</b>	<b>Quelques exemples d'application</b>	<b>25</b>
<b>4</b>	<b>Schémas numériques multi-pas</b>	<b>32</b>
4.1	Consistance, ordre, stabilité et convergence . . . . .	33
4.2	Schémas d'Adams-Baschforth . . . . .	35
4.2.1	Erreur de consistance et ordre de la méthodes $AB_{r+1}$ . . . . .	37
4.2.2	Stabilité de la méthode $AB_{r+1}$ . . . . .	38
4.3	Schémas d'Adams-Moulton . . . . .	39
4.3.1	Erreur de consistance et ordre de la méthode $AM_{r+1}$ . . . . .	41
4.3.2	Stabilité de la méthode $AM_{r+1}$ . . . . .	42
<b>5</b>	<b>Méthodes de prédiction-correction</b>	<b>43</b>
5.1	Exemples de méthodes de prédiction-correction en mode PECE . . . . .	44
5.2	Ordre et consistance des méthodes PECE . . . . .	45
5.3	Stabilité des méthodes PECE . . . . .	46
<b>6</b>	<b>Mise en œuvre</b>	<b>47</b>
6.1	La méthode d'Euler explicite . . . . .	47
6.2	La méthode d'Euler implicite . . . . .	48
6.3	La méthode de Runge Kutta 4 . . . . .	49
6.4	La méthode d'Adams Bashforth d'ordre 3 . . . . .	50
6.5	La méthode de prédiction correction . . . . .	51
6.6	La résolution numérique d'un système d'EDO : Modèle proie/prédateur . . . . .	52
<b>7</b>	<b>Exercices</b>	<b>54</b>
	<b>Annexe</b>	<b>57</b>



Nous abordons dans ce cours la résolution numérique du problème de Cauchy pour les équations différentielles ordinaires (EDO) du premier ordre. Après un bref rappel des notions de base sur les EDO, et de résultats développés dans le cours d'équations différentielles, nous introduisons deux classes de méthodes les plus couramment utilisées pour l'approximation numérique des équations scalaires : schémas à un pas et schémas multi-pas en présentant les concepts de consistance, stabilité, ordre et convergence qui permettent l'analyse théorique de ces méthodes. Divers exemples et exercices accompagnent le présent document afin d'assimiler les notions plus théoriques vues en cours.

# 1 Rappels sur le cours d'équations différentielles

## 1.1 Généralités

**Définition 1.1** Soit  $U$  un ouvert de  $\mathbb{R} \times \mathbb{R}^d$  et  $f : U \rightarrow \mathbb{R}^d$  une application supposée au moins continue par rapport aux deux variables. On appelle **solution** de l'équation différentielle

$$(E) \quad y'(t) = f(t, y(t)),$$

tout couple  $(J, y)$  où  $J \subset \mathbb{R}$  est un intervalle de  $\mathbb{R}$  et  $y$  une fonction dérivable définie sur  $J$  telle que

$$\forall t \in J, (t, y(t)) \in U \text{ et } y'(t) = f(t, y(t)).$$

**Définition 1.2** Toute solution  $(J, y)$  de  $(E)$  définie sur l'intervalle  $J = \mathbb{R}$  tout entier est dite **globale**.

Il arrive qu'on ne cherche pas toutes les solutions d'une EDO mais seulement celles qui vérifient certaines conditions, dites conditions initiales de Cauchy ou tout simplement conditions de Cauchy.

**Définition 1.3** Soit  $f : U \rightarrow \mathbb{R}^d$ ,  $(t_0, y_0) \in U$ . On appelle **solution du problème de Cauchy** associée à la donnée  $(t_0, y_0)$  toute solution  $(J, y)$  de l'équation  $(E)$  vérifiant de plus

$$(I) \quad t_0 \in J \text{ et } y(t_0) = y_0.$$

On obtient en intégrant  $\{(E), (I)\}$  entre  $t_0$  et  $t$

$$y(t) - y_0 = \int_{t_0}^t f(s, y(s)) ds. \tag{1.1}$$

La solution  $y$  du problème de Cauchy  $\{(E), (I)\}$  est de classe  $\mathcal{C}^1$  et satisfait l'équation intégrale (1.1). Inversement, si  $y$  est définie par (1.1), alors elle est continue sur  $J$  et  $y(t_0) = y_0$ . De plus, en tant que primitive de la fonction continue  $f(., y(.))$ ,  $y \in \mathcal{C}^1(J)$  et satisfait l'équation différentielle  $y'(t) = f(t, y(t))$ . Ainsi, si  $f$  est continue, le problème de Cauchy  $\{(E), (I)\}$  est équivalent à l'équation intégrale (1.1). Nous verrons plus loin comment tirer parti de cette équivalence pour les méthodes numériques.

**-Cylindre de sécurité :** On choisit un compact  $K_0 = [t_0 - T_0, t_0 + T_0] \times \overline{B}(y_0, r_0) \subset U$ , centré en  $(t_0, y_0)$ , où  $T_0, r_0 > 0$  et  $\overline{B}(y_0, r_0)$  est la boule fermée de centre  $y_0$  et de rayon  $r_0$ . On note  $M = \sup_{(t,y) \in K_0} \|f(t, y)\|$ . Ce nombre  $M$  est fini par la compacité de  $K_0$  et la continuité de  $f$ . le choix de la norme sur  $\mathbb{R}^d$  est arbitraire.

**Proposition 1.4** Soit  $T := \min(T_0, \frac{r_0}{M})$ . Toute solution  $y$  au problème de Cauchy pour les conditions initiales  $(t_0, y_0)$  satisfait à

$$|t - t_0| \leq T \implies \|y(t) - y_0\| \leq r_0.$$

Cette dernière implication veut dire aucune solution  $y$  ne peut "s'échapper" de  $[t_0 - T, t_0 + T] \times \overline{B}(y_0, r_0)$ . On dit dans ce cas que  $[t_0 - T, t_0 + T] \times \overline{B}(y_0, r_0)$  est un **cylindre de sécurité** pour l'équation (E).

## 1.2 Énoncé du théorème de Cauchy-Lipschitz

Nous renvoyons au cours d'équations différentielles pour la démonstration à l'aide du théorème du point fixe et pour d'éventuels énoncés plus généraux.

**Définition 1.5** Une fonction  $f : U \rightarrow \mathbb{R}^d$  est dite **localement lipschitzienne** par rapport à la variable d'état (ou à la seconde variable) si pour tout  $(t_0, y_0) \in U$ , il existe une constante  $C_{t_0, y_0} > 0$  et un voisinage  $V$  de  $(t_0, y_0)$  dans  $U$  tel que

$$\forall (t, y_1) \in V \text{ et } (t, y_2) \in V, \|f(t, y_1) - f(t, y_2)\| \leq C_{t_0, y_0} \|y_1 - y_2\|.$$

**Théorème 1.6 (Cauchy-Lipschitz)** On suppose que  $f : U \rightarrow \mathbb{R}^d$  est continue et localement lipschitzienne par rapport à la variable d'état, alors pour tout cylindre de sécurité  $C = [t_0 - T, t_0 + T] \times \overline{B}(y_0, r_0)$ , l'équation (E) **admet** une **unique** solution de condition initiale  $(t_0, y_0)$  définie sur l'intervalle  $[t_0 - T, t_0 + T]$ .

Un exemple d'équation pour laquelle le problème de Cauchy n'a pas de solution unique est l'équation

$$y'(t) = 2\sqrt{|y|},$$

qui possède une infinité de solutions vérifiant  $y(0) = 0$  dont la fonction identiquement nulle et la fonction  $y(t) = |t|t$ .

## 1.3 Sur la régularité des solutions

Rappelons qu'une fonction de plusieurs variables est dite de classe  $\mathcal{C}^k$  avec  $k$  un entier  $\geq 1$  si elle admet des dérivées partielles continues jusqu'à l'ordre  $k$ .

Si la fonction  $f$  est de classe  $\mathcal{C}^k$ , on peut en déduire un résultat sur l'ordre de dérivabilité des solutions  $y$  et calculer les dérivées successives de  $y$  à l'aide de fonctions construites récursivement à partir de  $f$ . C'est utile pour obtenir des majorations fondées sur la formule de Taylor.

**Proposition 1.7**

i) : Si la fonction  $f = \begin{pmatrix} f_1 \\ \vdots \\ f_d \end{pmatrix}$  est de classe  $\mathcal{C}^k, k \geq 1$ , toute solution  $t \mapsto y(t)$ , de l'équation

(E) est de classe  $\mathcal{C}^{k+1}$ .

ii) : Les fonctions  $f^{[i]} : U \rightarrow \mathbb{R}$ , définies récursivement pour  $i = 0, \dots, k$ , par les formules

$$f^{[0]} = f, \quad f^{[i+1]}(t, y) = \frac{\partial f^{[i]}}{\partial t}(t, y) + \sum_{\ell=0}^d f_\ell(t, y) \frac{\partial f^{[i]}}{\partial y_\ell}(t, y)$$

sont de classes respectives  $\mathcal{C}^{k-i}$  et pour toute solution  $y$  de l'équation (E) ses dérivées successives s'obtiennent par les formules

$$y^{(i+1)}(t) = f^{[i]}(t, y(t))$$

**Preuve :** On démontre par récurrence sur  $i$ .

-La classe de différentiabilité de  $f^{[i]}$ , s'obtient par la formule de récurrence et une application directe de la définition de la classe  $\mathcal{C}^k$ .

-Pour la différentiabilité de la solution  $y$ , on montre par récurrence sur  $i$ , pour  $0 \leq i \leq k$ , l'existence de  $y^{(i+1)}$  et la formule annoncée en ii) : pour  $i = 0$ , il s'agit de l'égalité  $y'(t) = f(t, y(t))$  qui signifie que  $t \mapsto y(t)$  est une solution. On voit sur cette relation que  $y'$  est continûment dérivable et la formule de dérivation des fonctions composées s'écrit :

$$\begin{aligned} y''(t) &= \frac{\partial f}{\partial t}(t, y(t)) + \sum_{\ell=0}^d y'_\ell(t) \frac{\partial f}{\partial y_\ell}(t, y(t)) \\ &= \frac{\partial f}{\partial t}(t, y(t)) + \sum_{\ell=0}^d f_\ell(t, y(t)) \frac{\partial f}{\partial y_\ell}(t, y(t)) \\ &= f^{[1]}(t, y(t)). \end{aligned}$$

Pour le pas général de la récurrence, on suppose que  $y^{(i+1)}(t) = f^{[i]}(t, y(t))$ , avec  $1 \leq i \leq k-1$ , alors puisque  $f^{[i]}$  est de classe  $\mathcal{C}^{k-i}$ , avec  $k-i \geq 1$ , on en déduit que  $y^{(i+1)}$  est encore continûment dérivable et par un calcul semblable au cas de  $y''$ , on trouve

$$\begin{aligned} y^{(i+2)}(t) = (y^{(i+1)})'(t) &= \frac{\partial f^{[i]}}{\partial t}(t, y(t)) + \sum_{\ell=0}^d f_\ell(t, y(t)) \frac{\partial f^{[i]}}{\partial y_\ell}(t, y(t)) \\ &= f^{[i+1]}(t, y(t)). \end{aligned}$$

■

**Exemple 1.8** Dans le cas scalaire ( $d = 1$ ), voici les formules obtenues pour  $f^{[1]}$  et  $f^{[2]}$  qui permettent de calculer  $y^{(2)}$  et  $y^{(3)}$  :

$$\begin{aligned} f^{[1]} &= \frac{\partial f}{\partial t} + f \frac{\partial f}{\partial y} \\ f^{[2]} &= \frac{\partial f^{[1]}}{\partial t} + f \frac{\partial f^{[1]}}{\partial y} \\ &= \frac{\partial^2 f}{\partial t^2} + 2f \frac{\partial^2 f}{\partial t \partial y} + f^2 \frac{\partial^2 f}{\partial y^2} + f \left( \frac{\partial f}{\partial y} \right)^2 + \frac{\partial f}{\partial t} \frac{\partial f}{\partial y}. \end{aligned}$$

## 2 Méthodes numériques à un pas

Pour des équations différentielles d'un intérêt pratique, on trouve rarement la solution  $y(t)$  exprimée avec une formule exacte. On a vu par exemple dans la section précédente qu'une solution de l'équation

$$y'(t) = f(t, y(t))$$

est obtenue en intégrant de  $t_0$  à  $t$  :

$$y(t) = y(t_0) + \int_{t_0}^t f(s, y(s)) ds.$$

Le problème avec cette solution est que l'inconnue  $y$  se trouve sous l'intégrale.

On reprend les notations précédentes concernant un cylindre de sécurité, et on ne s'occupe pour simplifier que des solutions à droite c'est à dire sur l'intervalle  $[t_0, t_0 + T]$ .

L'objectif de cette section est de décrire un certain nombre de méthodes permettant de résoudre numériquement le problème de Cauchy de condition initiale

$$(I) \quad y(t_0) = y_0$$

pour une équation différentielle

$$(E) \quad y'(t) = f(t, y(t)).$$

Dans ce cours, on supposera que  $f$  satisfait aux conditions du théorème de Cauchy-Lipschitz. Ceci assure que le problème  $\{(E), (I)\}$  admet une unique solution. On notera  $t \mapsto y(t)$  la solution unique du problème sur  $[t_0, t_0 + T]$  dont le graphe est contenu dans un cylindre de sécurité  $[t_0 - T, t_0 + T] \times \overline{B}(y_0, r_0)$ .

Nous avons choisi ici d'exposer le cas des équations unidimensionnelles dans le seul but de simplifier les notations ; le cas des systèmes dans  $\mathbb{R}^d$  est tout à fait identique, à condition de considérer  $y$  et  $f$  comme des fonctions vectorielles.

Etant donné une subdivision  $t_0 < t_1 < \dots < t_N = t_0 + T$  de  $[t_0, t_0 + T]$ , on cherche à déterminer des valeurs approchées  $y_n$  des valeurs  $y(t_n)$ ,  $0 \leq n \leq N$ , prises par la solution exacte  $y$ . On notera les pas successifs

$$h_n = t_{n+1} - t_n, \quad 0 \leq n \leq N - 1,$$

et

$$h_{\max} = \max_n(h_n)$$

le maximum du pas.

**Définition 2.1** Une méthode (ou schéma) à un pas **explicite** est une équation de récurrence de la forme

$$\begin{cases} y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n), & 0 \leq n \leq N - 1, \\ t_{n+1} = t_n + h_n \end{cases}$$

Le domaine de définition de  $\Phi$  contient au moins  $U \times [0, \delta]$ ,  $\delta > 0$ .

**Définition 2.2** Un schéma à un pas est dit **implicite** s'il est de la forme

$$\begin{cases} y_{n+1} = y_n + h_n \Phi(t_n, y_n, y_{n+1}, h_n), & 0 \leq n \leq N-1, \\ t_{n+1} = t_n + h_n \end{cases}$$

C'est-à-dire si  $\Phi$  dépend non linéairement de  $y_{n+1}$ .

Pour ce type de méthodes il s'agira le plus souvent de s'assurer que l'équation

$$y = y_n + h \Phi(t_n, y_n, y, h)$$

a une unique solution du moins pour tout  $h$  assez petit. Dans les cas les plus courants cela résultera du théorème des fonctions implicites.

## 2.1 Quelques exemples de méthodes à un pas

### 2.1.1 Méthodes d'Euler

Une façon d'obtenir une multitude de schémas, est d'intégrer l'EDO ( $E$ ) sur  $[t_n, t_{n+1}]$  :

$$y(t_{n+1}) - y(t_n) = \int_{t_n}^{t_{n+1}} f(t, y(t)) dt,$$

et ensuite d'approcher l'intégrale.

Par exemple

- Intégration par la méthode des rectangles à gauche

$$\int_{t_n}^{t_{n+1}} f(t, y(t)) dt \approx h_n f(t_n, y(t_n)),$$

ce qui donne le schéma d'**Euler explicite**

$$y_{n+1} = y_n + h_n f(t_n, y_n). \quad (2.1)$$

- Intégration par la méthode des rectangles à droite

$$\int_{t_n}^{t_{n+1}} f(t, y(t)) dt \approx h_n f(t_{n+1}, y(t_{n+1})),$$

ce qui donne le schéma d'**Euler implicite**

$$y_{n+1} = y_n + h_n f(t_{n+1}, y_{n+1}). \quad (2.2)$$

- Intégration par la méthode du point milieu

$$\int_{t_n}^{t_{n+1}} f(t, y(t)) dt \approx h_n f\left(t_n + \frac{h_n}{2}, y\left(t_n + \frac{h_n}{2}\right)\right), \quad (2.3)$$

ici, on connaît uniquement la valeur de  $y_n$ , et pour donner une approximation de la solution au point  $t_n + \frac{h_n}{2}$ , on utilise le schéma d'Euler explicite

$$y\left(t_n + \frac{h_n}{2}\right) \approx y(t_n) + \frac{h_n}{2} f(t_n, y(t_n)),$$

ce qui donne le schéma d'**Euler modifié**

$$y_{n+1} = y_n + h_n \left( f\left(t_n + \frac{h_n}{2}, y_n + \frac{h_n}{2} f(t_n, y_n)\right) \right).$$



**Exemple : Explicite ou implicite ?** On lit souvent que "les schémas implicites sont plus stables" (voir plus loin pour la définition de la notion de stabilité). Il est donc naturel de les préférer aux schémas explicites. Cependant, dans le cas général, le choix n'est pas si évident, comme nous allons le voir sur des exemples, en étudiant le comportement respectif des schémas d'Euler explicite et implicite.

• **Ici l'implicite gagne :** on considère le problème de Cauchy suivant

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

dans l'intervalle temps  $[0, T]$ . La solution exacte est clairement donnée par  $y(t) = e^{-t}$ . On suppose que les pas de temps sont constants, c'est-à-dire  $h_n = h = \frac{T}{N}, 0 \leq n \leq N - 1$ . Le schéma d'Euler explicite s'écrit dans ce cas :

$$y_{n+1} = y_n - hy_n = (1 - h)y_n,$$

et donc

$$y_{n+1} = (1 - h)^n, \tag{2.4}$$

d'où

$$y_N = (1 - h)^{\frac{T}{h}} \rightarrow e^{-T} \text{ quand } h \rightarrow 0.$$

Lorsqu'on cherche par exemple à obtenir le comportement de la solution d'une équation différentielle "dans les grands temps", on peut être amené à utiliser des pas de discrétisation relativement grands. Ceci peut être aussi le cas dans des problèmes de couplage avec d'autres équations, les "échelles de temps" des équations pouvant être très différentes pour les différentes équations. Que se passe-t-il dans ce cas ? Dans le cas de notre exemple, si on prend  $h = 2$ , on obtient alors  $y_n = (-1)^n$ , ce qui n'est clairement pas une bonne approximation de la solution. Dans un problème d'origine physique où  $y$  serait une concentration ou une densité, il est indispensable que le schéma respecte la positivité. Dans l'exemple présent, le schéma d'Euler explicite (2.4) ne donne pas une solution approchée raisonnable pour  $h$  grand.

Si on essaye maintenant de calculer une solution approchée à l'aide du schéma d'Euler implicite, on obtient

$$y_{n+1} = y_n - hy_{n+1},$$

et donc

$$y_n = \frac{1}{(1 + h)^n}, 0 \leq n \leq N.$$

Dans ce cas, la solution approchée reste "proche" de la solution exacte, et positive, même pour des pas de discrétisation grands. On pourrait en conclure un peu hâtivement que le schéma implicite est "meilleur" que le schéma explicite. On va voir dans l'exemple qui suit qu'une telle conclusion est peu rapide.

• **Ici l'implicite perd :** On considère maintenant le problème de Cauchy

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

La solution exacte est maintenant  $y(t) = e^t$ . On prend toujours un pas de discrétisation constant égal à  $h$ , le schéma d'Euler explicite s'écrit :

$$y_{n+1} = y_n + hy_n = (1 + h)y_n,$$

et donc

$$y_{n+1} = (1 + h)^n, 0 \leq n \leq N,$$

d'où

$$y_N = (1 + h)^{\frac{T}{h}} \rightarrow e^T \text{ quand } h \rightarrow 0.$$

Contrairement à l'exemple précédent, la solution approchée donnée par le schéma d'Euler explicite reste "raisonnable" même pour les grands pas de temps. Si on essaye maintenant de calculer une solution approchée à l'aide du schéma d'Euler implicite, on obtient

$$y_{n+1} = y_n + hy_{n+1}$$

c.à.d

$$y_{n+1} = \frac{1}{1 - h} y_n$$

On remarque d'une part que le schéma implicite n'est pas défini pour  $h = 1$ , et que d'autre part si  $h$  est proche de 1 (par valeurs supérieures ou inférieures), la solution approchée "explose". De plus pour les valeurs de  $h$  supérieures à 1, on perd la positivité de la solution (pour  $h = 2$  par exemple la solution approchée oscille entre les valeurs  $+1$  et  $-1$ ). Dans le cadre de cet exemple, le choix explicite semble donc plus approprié.

En conclusion de ces deux exemples, il semble que le "meilleur" schéma n'existe pas dans l'absolu. Le schéma de discrétisation doit être choisi en fonction du problème; ceci nécessite une bonne compréhension du comportement des schémas en fonction des problèmes donnés, donc une certaine expérience.

### 2.1.2 Méthode de Taylor d'ordre $p$

Supposons maintenant que  $f$  soit de classe  $\mathcal{C}^p$ , on a vu alors que la solution exacte  $y$  de (E) est de classe  $\mathcal{C}^{p+1}$  et on a défini des fonctions  $f^{[k]}$  construites par récurrence à partir de  $f$  et de ses dérivées partielles telles que  $y^{(k)}(t) = f^{[k-1]}(t, y(t))$ , pour  $k = 1, \dots, p + 1$ . La formule de Taylor d'ordre  $p$  donne

$$y(t_n + h_n) = y(t_n) + \sum_{k=1}^p \frac{1}{k!} h_n^k f^{[k-1]}(t_n, y(t_n)) + o(h_n^p),$$

ou avec la formule de Taylor avec reste de Lagrange :

$$y(t_n + h_n) = y(t_n) + \sum_{k=1}^p \frac{1}{k!} h_n^k f^{[k-1]}(t_n, y(t_n)) + \frac{1}{(p+1)!} h_n^{p+1} f^{[p]}(t_n + \theta h_n, y(t_n + \theta h_n)), \quad \theta \in ]0, 1[. \quad (2.5)$$

On est donc amené à considérer l'algorithme suivant, appelé **méthode de Taylor** d'ordre  $p$  :

$$(\mathcal{T}_p) \quad \begin{cases} y_{n+1} = y_n + \sum_{k=1}^p \frac{1}{k!} h_n^k f^{[k-1]}(t_n, y_n), \\ t_{n+1} = t_n + h_n. \end{cases}$$

D'après la définition 2.1, cet algorithme correspond au choix  $\Phi(t, y, h) = \sum_{k=1}^p \frac{1}{k!} h^{k-1} f^{[k-1]}(t, y(t))$ . On peut remarquer facilement que la méthode d'Euler n'est autre que la méthode de Taylor ( $\mathcal{T}_1$ ).

**Remarque 2.3** *La méthode de Taylor n'est en général pas utilisée en pratique car le calcul des valeurs  $f^{[k]}$  est trop coûteux.*

## 2.2 Etude des méthodes à un pas

Dans la section précédente, plusieurs schémas numériques ont été présentés et afin de les comparer plusieurs critères seront étudiés. Les trois principaux sont la convergence, la stabilité et la consistance, permettant de relier la solution exacte à la solution approchée. L'étude concernera les méthodes explicites à un pas.

**Définition 2.4** (*Consistance*) *Si on remplace la solution approchée par la solution exacte dans le schéma numérique, quelle est l'erreur commise en fonction de  $h_n$ ? Pour ceci, on définit l'erreur locale de consistance*

$$e_n = y(t_{n+1}) - \bar{y}_{n+1}, \quad (2.6)$$

avec  $\bar{y}_{n+1}$  est la solution du schéma issue de  $y(t_n)$ , i.e.

$$\bar{y}_{n+1} = y(t_n) + h_n \Phi(t_n, y(t_n), h_n).$$

**Définition 2.5** *On dit que la méthode est consistante si pour toute solution exacte  $y$  la somme des erreurs de consistance relatives à  $y$ , soit  $\sum_{n=0}^{N-1} |e_n|$ , tend vers 0 quand  $h_{\max}$  tend vers 0.*

Il existe une méthode simple pour vérifier la consistance des méthodes numériques. Dans l'énoncé suivant, on suppose que  $\Phi$  remplit la condition suivante presque toujours réalisée dans les exemples usuels :  $\Phi$  est continue sur un ouvert contenant  $U \times [0, \delta]$ ,  $\delta > 0$ .

**Théorème 2.6** (*Condition nécessaire et suffisante de consistance*)

*La méthode à un pas définie par la fonction  $\Phi$  est consistante si et seulement si*

$$\forall (t, y) \in U, \quad \Phi(t, y, 0) = f(t, y).$$

**Preuve :** Soit  $y$  une solution exacte de l'équation (E) et soient

$$e_n = y(t_{n+1}) - y(t_n) - h_n \Phi(t_n, y(t_n), h_n)$$

les erreurs de consistance correspondantes. D'après le théorème des accroissements finis, il existe  $c_n \in ]t_n, t_{n+1}[$  tel que

$$y(t_{n+1}) - y(t_n) = h_n y'(c_n) = h_n f(c_n, y(c_n)),$$

d'où

$$e_n = h_n (f(c_n, y(c_n)) - \Phi(t_n, y(t_n), h_n)) = h_n (\alpha_n + \beta_n),$$

avec

$$\alpha_n = f(c_n, y(c_n)) - \Phi(c_n, y(c_n), 0),$$

$$\beta_n = \Phi(c_n, y(c_n), 0) - \Phi(t_n, y(t_n), h_n).$$

D'après l'uniforme continuité de  $\Phi$  sur  $C \times [0, \delta]$ , où  $C$  est le cylindre de sécurité sur lequel on travaille : on a

$$\forall \varepsilon > 0, \exists \eta > 0, \exists \gamma > 0, \text{ tels que } |h| < \eta, |t_1 - t_2| < \eta, |y_1 - y_2| < \gamma \Rightarrow |\Phi(t_1, y_1, 0) - \Phi(t_2, y_2, h)| < \varepsilon.$$

Par ailleurs, quitte à diminuer  $\eta$ , on peut s'assurer en utilisant l'uniforme continuité de  $y$  sur  $[t_0, t_0 + T]$  que

$$h_{\max} < \eta \Rightarrow |y(c_n) - y(t_n)| < \gamma.$$

En enchaînant les deux implications précédentes

$$h_{\max} < \eta \Rightarrow \Phi(c_n, y(c_n), 0) - \Phi(t_n, y(t_n), h_n) < \varepsilon.$$

Par conséquent

$$h_{\max} < \eta \Rightarrow \sum_{n=0}^{N-1} h_n |\Phi(c_n, y(c_n), 0) - \Phi(t_n, y(t_n), h_n)| < \varepsilon \sum_{n=0}^{N-1} h_n = \varepsilon T.$$

avec les notations abrégées, on obtient

$$\left| \sum_{n=0}^{N-1} |e_n| - \sum_{n=0}^{N-1} h_n |\alpha_n| \right| \leq \sum_{n=0}^{N-1} h_n |\beta_n| \leq \varepsilon \sum_{n=0}^{N-1} h_n = \varepsilon T.$$

On en déduit

$$\begin{aligned} \lim_{h_{\max} \rightarrow 0} \sum_{n=0}^{N-1} |e_n| &= \lim_{h_{\max} \rightarrow 0} \sum_{n=0}^{N-1} h_n |\alpha_n| \\ &= \int_{t_0}^{t_0+T} |f(t, y(t)) - \Phi(t, y(t), 0)| dt, \end{aligned}$$

car  $\sum_{n=0}^{N-1} h_n |\alpha_n|$  est une somme de Riemann de l'intégrale précédente.

La condition de consistance est équivalente au fait que cette limite est nulle donc à

$$\int_{t_0}^{t_0+T} |f(t, y(t)) - \Phi(t, y(t), 0)| dt = 0.$$

La nullité de cette intégrale de fonction positive continue impose pour tout  $t$ ,  $f(t, y(t)) = \Phi(t, y(t), 0)$ . Dans tout ce raisonnement la condition initiale  $(t_0, y_0)$  est arbitraire, et donc l'égalité  $f(t_0, y_0) = \Phi(t_0, y_0, 0)$  est valable pour tout  $(t_0, y_0) \in U$ . ■

**Définition 2.7** On dit qu'une méthode à un pas est d'ordre  $\geq p$  si pour toute solution exacte  $y$  de  $(E)$ , il existe une constante  $C \geq 0$  telle que l'erreur de consistance vérifie

$$|e_n| \leq C h_n^{p+1}, \quad \forall n, \quad 0 \leq n \leq N.$$

Elle est dite d'ordre  $p$  (exactement) si elle est d'ordre  $\geq p$  mais pas d'ordre  $\geq p+1$ .

**Remarque 2.8** Des définitions 2.5 et 2.7, on déduit immédiatement qu'une méthode à un pas est consistante si elle est au moins d'ordre 1.

**Exemple 2.9** La méthode de Taylor ( $\mathcal{T}_p$ ) est du point de vue de l'erreur de consistance d'ordre  $p$ . Plus précisément si on considère un cylindre de sécurité  $C = [t_0 - T, t_0 + T] \times \overline{B}(y_0, r)$ , on a la majoration

$$e_n \leq \frac{1}{(p+1)!} \sup_{(t,y) \in C} \|f^{[p]}(t, y)\| h_n^{p+1}.$$

En effet, selon la formule de Taylor qui a servi de base à la méthode, on obtient directement en prenant la formule avec reste de Lagrange (2.5)

$$\begin{aligned} e_n &= y(t_n + h_n) - y(t_n) - \sum_{k=1}^p h_n^k \frac{1}{k!} f^{[k-1]}(t_n, y(t_n)) \\ &= \frac{1}{(p+1)!} f^{[p]}(t_n + \theta h_n, y(t_n + \theta h_n)) h_n^{p+1} \leq \frac{1}{(p+1)!} \sup_{(t,y) \in C} \|f^{[p]}(t, y)\| h_n^{p+1}. \end{aligned}$$

Décrivons maintenant une méthode générale permettant de calculer l'ordre de consistance : l'erreur de consistance est donnée par

$$e_n = y(t_{n+1}) - y(t_n) - h_n \Phi(t_n, y(t_n), h_n).$$

Si on suppose que  $\Phi$  est de classe  $\mathcal{C}^p$ , et puisque  $f$  est de classe  $\mathcal{C}^p$ , la solution  $y$  est de classe  $\mathcal{C}^{p+1}$ , par conséquent, en appliquant la formule de Taylor avec reste de Lagrange (2.5), on a l'existence de  $c_n, d_n \in ]t_n, t_{n+1}[$  tels que :

$$\begin{aligned} y(t_{n+1}) - y(t_n) &= y(t_n + h_n) - y(t_n) \\ &= \sum_{k=1}^p \frac{1}{k!} h_n^k y^{(k)}(t_n) + \frac{1}{(p+1)!} h_n^{p+1} y^{(p+1)}(c_n) \\ &= \sum_{\ell=0}^{p-1} \frac{1}{(k+1)!} h_n^{k+1} f^{[k]}(t_n, y(t_n)) + \frac{1}{(p+1)!} h_n^{p+1} f^{[p]}(c_n, y(c_n)), \end{aligned}$$

et

$$\Phi(t_n, y(t_n), h_n) = \sum_{\ell=0}^{p-1} \frac{1}{\ell!} h_n^\ell \frac{\partial^\ell \Phi}{\partial h^\ell}(t_n, y(t_n), 0) + \frac{h_n^p}{p!} \frac{\partial^p \Phi}{\partial h^p}(t_n, y(t_n), d_n).$$

On en tire

$$e_n = h_n \left( \sum_{\ell=0}^{p-1} \frac{h_n^\ell}{\ell!} \left( \frac{f^{[\ell]}(t_n, y(t_n))}{\ell+1} - \frac{\partial^\ell \Phi(t_n, y(t_n), 0)}{\partial h^\ell} \right) \right) + \frac{h_n^{p+1}}{p!} \left( \frac{f^{[p]}(c_n, y(c_n))}{p+1} - \frac{\partial^p \Phi(t_n, y(t_n), d_n)}{\partial h^p} \right).$$

Pour que cette expression soit un développement limité en  $h_n$  de la forme  $o(h_n^p)$ , il faut et il suffit que tous les termes  $\frac{f^{[\ell]}(t_n, y(t_n))}{\ell+1} - \frac{\partial^\ell \Phi(t_n, y(t_n), 0)}{\partial h^\ell}$  soient nuls. Le reste fournit la majoration de l'énoncé avec la constante

$$C = \frac{1}{(p+1)!} \|f^{[p]}\|_\infty + \frac{1}{p!} \left\| \frac{\partial^p \Phi}{\partial h^p} \right\|_\infty,$$

où les normes utilisées sont les normes du sup  $\|\cdot\|_\infty$  et les solutions sont supposées confinées dans le compact  $[t_0, t_0 + T] \times J \times [0, \delta]$ , où  $J$  est un compact.

On dispose alors du résultat suivant

**Proposition 2.10** *Sous l'hypothèse que  $f$  et  $\Phi$  sont de classe  $\mathcal{C}^p$ , une méthode à un pas est d'ordre  $p$  si et seulement si les conditions suivantes sont remplies :*

$$\frac{\partial^\ell \Phi}{\partial h^\ell}(t_n, y(t_n), 0) = \frac{1}{(\ell + 1)!} f^{[\ell]}(t_n, y(t_n)), \text{ pour } 0 \leq \ell \leq p - 1, \quad 0 \leq n \leq N.$$

Dans la pratique les valeurs de  $y_n$  sont perturbées par des valeurs voisines  $\tilde{y}_n$  pour deux raisons :

- 1) *Erreurs d'arrondis* : on représente en machine la valeur  $y_n$  issue du calcul par un nombre décimal à  $q$  chiffres.  $|y_n - \tilde{y}_n|$  est alors l'erreur d'arrondi majorée en valeur relative par  $10^{-q} y_n$ .
- 2) *Incertitude expérimentale* : dans la plupart des problèmes concrets, la "vraie" valeur de  $y_0$  est remplacée par une valeur  $\tilde{y}_0$  tirée d'une expérience, d'une hypothèse, etc,  $|y_0 - \tilde{y}_0|$  est donc majorée par un nombre qui dépend de la précision expérimentale.

La méthode ne peut donc être utile que si la perturbation sur  $y_n - \tilde{y}_n$  provoquée par une faible perturbation  $|y_0 - \tilde{y}_0|$  des données initiales et par les erreurs d'arrondi est faible. On est donc amené à la définition suivante.

**Définition 2.11** *On dit que la méthode est stable s'il existe une constante  $S \geq 0$ , appelée constante de stabilité, telle que pour toutes suites  $(y_n), (\tilde{y}_n)$  définies par*

$$y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n), \quad 0 \leq n < N,$$

$$\tilde{y}_{n+1} = \tilde{y}_n + h_n \Phi(t_n, \tilde{y}_n, h_n) + \varepsilon_n, \quad 0 \leq n < N,$$

on ait

$$\max_{0 \leq n \leq N} |\tilde{y}_n - y_n| \leq S \left( |\tilde{y}_0 - y_0| + \sum_{0 \leq n < N} |\varepsilon_n| \right).$$

où  $\varepsilon_n \in \mathbb{R}$  (appelé la perturbation).

Autrement dit, une petite erreur initiale  $|\tilde{y}_0 - y_0|$  et de petites erreurs d'arrondi  $\varepsilon_n$  dans le calcul récurrent des  $\tilde{y}_n$  provoquent une erreur finale  $\max_n |\tilde{y}_n - y_n|$  contrôlable.

**Théorème 2.12** *(Une condition suffisante de stabilité)*

*Si  $\Phi$  est lipschitzienne par rapport à la variable  $y$ , la méthode est stable. De plus si  $L$  est la constante de Lipschitz pour  $\Phi$ , la constante de stabilité est  $S = e^{LT}$ .*

La démonstration de ce théorème repose sur le lemme suivant :

**Lemme 2.13** *(de Gronwall discret) - Soient les suites  $h_n, \theta_n \geq 0$  et  $\varepsilon_n \in \mathbb{R}$  telles que*

$$\theta_{n+1} \leq (1 + Lh_n)\theta_n + |\varepsilon_n|,$$

alors

$$\theta_n \leq e^{L(t_n - t_0)} \theta_0 + \sum_{i=0}^{n-1} e^{L(t_n - t_{i+1})} |\varepsilon_i|.$$

**Preuve :** Le lemme se vérifie par récurrence sur  $n$ . Pour  $n = 0$ , l'inégalité se réduit à  $\theta_0 \leq \theta_0$ . Supposons maintenant l'inégalité vraie à l'ordre  $n$ . On sait que

$$1 + Lh_n \leq e^{Lh_n} = e^{L(t_{n+1} - t_n)}.$$

Par hypothèse on a

$$\begin{aligned}\theta_{n+1} &\leq (1 + Lh_n)\theta_n + |\varepsilon_n| \\ &\leq e^{L(t_{n+1}-t_n)}\theta_n + |\varepsilon_n| \\ &\leq e^{L(t_{n+1}-t_0)}\theta_0 + \sum_{i=0}^{n-1} e^{L(t_{n+1}-t_{i+1})}|\varepsilon_i| + |\varepsilon_n|.\end{aligned}$$

L'inégalité s'ensuit à l'ordre  $n + 1$ . ■

**Preuve du théorème 2.12 :** Considérons deux suites  $(y_n)$  et  $(\tilde{y}_n)$  telles que

$$\begin{aligned}y_{n+1} &= y_n + h_n\Phi(t_n, y_n, h_n), \\ \tilde{y}_{n+1} &= \tilde{y}_n + h_n\Phi(t_n, \tilde{y}_n, h_n) + \varepsilon_n.\end{aligned}$$

Par différence, on obtient

$$|\tilde{y}_{n+1} - y_{n+1}| \leq |\tilde{y}_n - y_n| + h_n|\Phi(t_n, y_n, h_n) - \Phi(t_n, \tilde{y}_n, h_n)| + |\varepsilon_n|.$$

On pose  $\theta_n = |y_n - \tilde{y}_n|$ . Par définition de la constante de Lipschitz pour  $\Phi$ ,

$$\Phi(t, y_1, h) - \Phi(t, y_2, h) \leq L|y_1 - y_2|,$$

quel que soient  $(t, y_1, h), (t, y_2, h)$  dans le domaine de définition de  $\Phi$ . La majoration suivante découle aussitôt de la définition de la suite  $(\theta_n)$  et de la condition de Lipschitz

$$\theta_{n+1} \leq (1 + h_n L)\theta_n + \varepsilon_n.$$

Comme  $t_n - t_0 \leq T$  et  $t_n - t_{i+1} \leq T$ , le lemme de Gronwall implique

$$\max_{0 \leq n \leq N} \theta_n \leq e^{LT} \left( \theta_0 + \sum_{i=0}^{N-1} |\varepsilon_i| \right),$$

et le théorème est démontré. ■

**Remarque 2.14** Ces calculs ne sont corrects que tant que les  $(t_n, y_n, h_n)$  et  $(t_n, \tilde{y}_n, h_n)$  restent dans le domaine où  $\Phi$  est Lipschitzienne de constante  $L$ .

Une autre notion très importante en pratique est la suivante :

**Définition 2.15** La convergence est une propriété de la solution numérique. Une méthode numérique est dite convergente si pour toute solution exacte  $y$  définie sur un intervalle  $[t_0, t_0 + T]$  et toute suite  $(y_n)$  construite selon le schéma numérique considéré, à partir de  $y_0$  et d'une subdivision de  $[t_0, t_0 + T]$  on a la relation de convergence uniforme :

$$\lim_{\substack{h_{\max} \rightarrow 0 \\ y_0 \rightarrow y(t_0)}} \max_{0 \leq n \leq N} |y_n - y(t_n)| \rightarrow 0.$$

La quantité  $\max_{0 \leq n \leq N} |y_n - y(t_n)|$  s'appelle l'erreur globale. C'est évidemment cette erreur qui importe dans la pratique.

La consistance et la stabilité sont en général relativement facile à démontrer. La convergence demande souvent des démonstrations longues et ardues. Le théorème suivant permet d'obvier à cette difficulté.

**Théorème 2.16** *Une méthode numérique à un pas qui est stable et consistante est convergente.*

**Preuve :** Posons  $\tilde{y}_n = y(t_n)$ . Dans ce cas l'erreur de consistance est par définition le réel  $e_n$  qui complète la formule

$$y(t_{n+1}) = \tilde{y}_{n+1} = \tilde{y}_n + h_n \Phi(t_n, \tilde{y}_n, h_n) + e_n.$$

Donc  $e_n$  joue pour la suite des  $y(t_n)$  le rôle de la perturbation  $\varepsilon_n$  introduite dans la définition 2.11. D'après la définition de la stabilité on a donc

$$\max_{0 \leq n \leq N} |y_n - y(t_n)| \leq S(|y(t_0) - y_0| + \sum_{n=0}^{N-1} e_n).$$

L'hypothèse de consistance donne alors immédiatement le résultat annoncé. ■

**Exemple 2.17** *Si  $f$  est lipschitzienne en  $y$ , les méthodes d'Euler explicite et du point milieu sont convergentes.*

En effet : d'après le Théorème 2.16, il suffit pour cela d'établir la consistance et la stabilité.

- La consistance est une conséquence directe du Théorème 2.6 et de l'égalité  $\Phi|_{h=0} = f$ . C'est aussi valable pour la méthode de Taylor d'ordre  $p$ .

- La stabilité se déduit du Théorème 2.12 et du fait que  $\Phi$  est lipschitzienne : pour la méthode d'Euler explicite c'est immédiat car  $\Phi = f$ . Pour la méthode du point milieu cela résulte des calculs suivants : on a

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

donc si  $f$  est lipschitzienne en  $y$  de constante de Lipschitz  $L$  on obtient :

$$\begin{aligned} |\Phi(t, y_1, h) - \Phi(t, y_2, h)| &\leq L|y_1 - y_2 + \frac{h}{2}(f(t, y_1) - f(t, y_2))| \\ &\leq L|y_1 - y_2| + \frac{hL}{2}|f(t, y_1) - f(t, y_2)| \\ &\leq (L + \frac{h_{\max}}{2}L^2)|y_1 - y_2|, \end{aligned}$$

d'où le caractère lipschitzien de  $\Phi$  avec la constante de Lipschitz  $\Lambda = L + \frac{h_{\max}}{2}L^2$ .

## 2.3 Influence de l'ordre de la méthode sur l'erreur globale

On considère une méthode consistante et stable de constante de stabilité  $S$  qui est d'ordre  $p$ . On a les majoration d'erreurs suivantes. D'abord le cumul des erreurs de consistance (qui n'est pas l'erreur globale!!) est :

$$\sum_{0 \leq n < N} e_n \leq C \sum_{0 \leq n < N} h_n^{p+1} \leq Ch_{\max}^p \sum_{0 \leq n \leq N} h_n = CTh_{\max}^p.$$



Par définition de la stabilité, on trouve alors :

$$\max_{0 \leq n < N} |y_n - y(t_n)| \leq S(|y_0 - y(t_0)| + CT h_{\max}^p).$$

L'erreur globale est donc majorée dans le cas d'absence (ou de négligence) d'erreur sur la condition initiale par

$$SCT h_{\max}^p.$$

Si la constante  $SCT$  n'est pas trop grande (disons  $\leq 10^2$ ), une méthode d'ordre 3 avec un pas maximum  $h_{\max} = 10^{-2}$  permet d'atteindre une précision globale de l'ordre de  $10^{-4}$ .

## 2.4 Méthodes de type Runge-Kutta

Les méthodes de type Runge-Kutta permettent d'obtenir une plus grande précision que les méthodes d'Euler (dans le sens où elles donnent en général des solutions numériques plus proches des solutions analytiques que les méthodes d'Euler). Cette précision est obtenue par l'utilisation d'un pas de calcul intermédiaire. Les deux méthodes de Runge-Kutta les plus employées sont l'algorithme dit ( $RK2$ ) à deux pas de calcul et l'algorithme dit ( $RK4$ ) à quatre pas de calcul.

### 2.4.1 Description de la méthode

On considère comme d'habitude le problème de Cauchy

$$\begin{cases} y'(t) = f(t, y(t)) & (E) \\ y(t_0) = y_0 & (I) \end{cases}$$

avec une solution exacte  $y(t)$  sur  $[t_0, t_0 + T]$  et une subdivision  $t_0 < t_1 < \dots < t_N = t_0 + T$ .

L'idée est de calculer par récurrence les points  $(t_n, y_n)$  en utilisant des points intermédiaires  $(t_{n,i}, y_{n,i})$  avec

$$t_{n,i} = t_n + c_i h_n, \quad 1 \leq i \leq q, \quad c_i \in [0, 1].$$

A chacun de ces points on associe la pente correspondante

$$p_{n,i} = f(t_{n,i}, y_{n,i}).$$

Soit  $y$  la solution exacte de l'équation ( $E$ ). On a

$$\begin{aligned} y(t_{n,i}) &= y(t_n) + \int_{t_n}^{t_{n,i}} f(t, y(t)) dt \\ &= y(t_n) + h_n \int_0^{c_i} f(t_n + u h_n, y(t_n + u h_n)) du, \end{aligned}$$

grâce au changement de variable  $t = t_n + u h_n$ . De même

$$y(t_{n+1}) = y(t_n) + h_n \int_0^1 f(t_n + u h_n, y(t_n + u h_n)) du$$

On se donne alors pour chaque  $i = 1, 2, \dots, q$  une méthode d'intégration approchée

$$\int_0^{c_i} g(t) dt \simeq \sum_{j=1}^{i-1} a_{ij} g(c_j), \quad (M_i)$$

ces méthodes pouvant être différentes. On se donne également une méthode d'intégration approchée sur  $[0, 1]$

$$\int_0^1 g(t)dt \simeq \sum_{j=1}^q b_j g(c_j). \quad (M)$$

En appliquant ces méthodes d'intégration à  $g(u) = f(t_n + uh_n, y(t_n + uh_n))$ , il vient

$$y(t_{n,i}) \simeq y(t_n) + h_n \sum_{j=1}^{i-1} a_{ij} f(t_{n,j}, y(t_{n,j})),$$

$$y(t_{n+1}) \simeq y(t_n) + h_n \sum_{j=1}^q b_j f(t_{n,j}, y(t_{n,j})).$$

La méthode de Runge-Kutta correspondante est définie par l'algorithme

$$\begin{cases} t_{n,i} = t_n + c_i h_n \\ y_{n,i} = y_n + h_n \sum_{j=1}^{i-1} a_{ij} p_{n,j} \\ p_{n,i} = f(t_{n,i}, y_{n,i}), \quad 1 \leq i \leq q \\ t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n \sum_{j=1}^q b_j p_{n,j} \end{cases}$$

Conventionnellement elle est représentée par le tableau

$(M_1)$	$c_1$	0	0	$\cdots$	0	0
$(M_2)$	$c_2$	$a_{21}$	0	$\cdots$	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$\vdots$	$\vdots$	$\vdots$		0	0
$(M_q)$	$c_q$	$a_{q1}$	$a_{q2}$	$\cdots$	$a_{qq-1}$	0
$(M)$		$b_1$	$b_2$	$\cdots$	$b_{q-1}$	$b_q$

où les méthodes d'intégration approchées correspondent aux lignes. On pose par convention  $a_{ij} = 0$  pour  $j \geq i$ .

**Hypothèse :** On supposera que les méthodes d'intégration  $(M_i)$  et  $(M)$  sont d'ordre 0 (voir annexe) *i.e.*

$$\sum_{j=1}^{i-1} a_{i,j} = c_i, \quad \sum_{j=1}^q b_j = 1. \quad (2.7)$$

En particulier, on aura toujours

$$c_1 = 0, \quad t_{n,1} = t_n, \quad y_{n,1} = y_n, \quad p_{n,1} = f(t_n, y_n).$$

## 2.4.2 Exemples

**Exemple 1 :** Pour  $q = 1$ , le seul choix possible est

0	0
	1

On a ici  $c_1 = 0$ ,  $a_{11} = 0$ ,  $b_1 = 1$ . L'algorithme est donné par

$$\begin{cases} p_{n,1} = f(t_n, y_n) \\ t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n p_{n,1} \end{cases}$$

Il s'agit de la méthode d'Euler explicite.

**Exemple 2 :** Pour  $q = 2$ , on considère les tableaux de la forme

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \alpha & \alpha & 0 \\ \hline & 1 - \frac{1}{2\alpha} & \frac{1}{2\alpha} \end{array}$$

où  $\alpha \in ]0, 1]$ .

L'algorithme s'écrit ici

$$\begin{cases} p_{n,1} = f(t_n, y_n) \\ t_{n,2} = t_n + \alpha h_n \\ y_{n,2} = y_n + \alpha h_n p_{n,1} \\ p_{n,2} = f(t_{n,2}, y_{n,2}) \\ t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n \left( \left(1 - \frac{1}{2\alpha}\right) p_{n,1} + \frac{1}{2\alpha} p_{n,2} \right), \end{cases}$$

ou encore sous forme condensée :

$$y_{n+1} = y_n + h_n \left( \left(1 - \frac{1}{2\alpha}\right) f(t_n, y_n) + \frac{1}{2\alpha} f(t_n + \alpha h_n, y_n + \alpha h_n f(t_n, y_n)) \right).$$

Pour  $\alpha = \frac{1}{2}$  par exemple, on retrouve la méthode du point milieu

$$y_{n+1} = y_n + h_n f\left(t_n + \frac{h_n}{2}, y_n + \frac{h_n}{2} f(t_n, y_n)\right).$$

**Exemple 3 :** Méthode de Runge-Kutta classique *i.e.* pour  $q = 4$  :

il s'agit de la méthode définie par le tableau

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} \end{array}$$

L'algorithme correspondant s'écrit

$$\left\{ \begin{array}{l} p_{n,1} = f(t_n, y_n) \\ t_{n,2} = t_n + \frac{1}{2}h_n \\ y_{n,2} = y_n + \frac{1}{2}h_n p_{n,1} \\ p_{n,2} = f(t_{n,2}, y_{n,2}) \\ y_{n,3} = y_n + \frac{1}{2}h_n p_{n,2} \\ p_{n,3} = f(t_{n,2}, y_{n,3}) \\ t_{n+1} = t_n + h_n \\ y_{n,4} = y_n + h_n p_{n,3} \\ p_{n,4} = f(t_{n+1}, y_{n,4}) \\ y_{n+1} = y_n + h_n \left( \frac{1}{6}p_{n,1} + \frac{2}{6}p_{n,2} + \frac{2}{6}p_{n,3} + \frac{1}{6}p_{n,4} \right) \end{array} \right.$$

On verra plus loin que cette méthode est d'ordre 4. Dans ce cas les méthodes d'intégration ( $M_i$ ) et ( $M$ ) utilisées sont respectivement :

$$(M_2) \quad \int_0^{\frac{1}{2}} g(t)dt \simeq \frac{1}{2}g(0) \quad \text{méthodes des rectangles à gauche,}$$

$$(M_3) \quad \int_0^{\frac{1}{2}} g(t)dt \simeq \frac{1}{2}g\left(\frac{1}{2}\right) \quad \text{méthodes des rectangles à droite,}$$

$$(M_4) \quad \int_0^1 g(t)dt \simeq g\left(\frac{1}{2}\right) \quad \text{méthodes du point milieu,}$$

$$(M) \quad \int_0^1 g(t)dt \simeq \frac{1}{6}g(0) + \frac{2}{6}g\left(\frac{1}{2}\right) + \frac{2}{6}g\left(\frac{1}{2}\right) + \frac{1}{6}g(1) \quad \text{méthode de Simpson.}$$

### 2.4.3 Stabilité des méthodes de Runge-Kutta

Les méthodes de Runge-Kutta sont des méthodes à un pas

$$y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n),$$

avec  $\Phi(t_n, y_n, h_n) = \sum_{1 \leq j \leq q} b_j p_{n,j}$ . La fonction  $\Phi$  est définie de manière explicite par

$$\left\{ \begin{array}{l} \Phi(t, y, h) = \sum_{1 \leq j \leq q} b_j f(t + c_j h, y_j) \quad \text{avec} \\ y_i = y + h \sum_{1 \leq j < i} a_{ij} f(t + c_j h, y_j), \quad 1 \leq i \leq q. \end{array} \right. \quad (2.8)$$

Supposons que  $f$  soit  $k$ -lipschitzienne en  $y$  (uniformément par rapport à  $t \in [t_0, t_0 + T]$  et  $h \in [0, h_{\max}]$ ). On va montrer que  $\Phi$  est alors également lipschitzienne. Soit  $z \in \mathbb{R}$  et supposons  $\Phi(t, z, h)$  et  $z_i$  définis à partir de  $z$  comme dans la formule (2.8).

**Lemme 2.18** Soit  $\alpha = \max_i \left( \sum_{1 \leq j \leq i} |a_{ij}| \right)$ . Alors

$$|y_i - z_i| \leq (1 + (\alpha k h) + (\alpha k h)^2 + \dots + (\alpha k h)^{i-1}) |y - z|.$$

**Preuve :** Le lemme se démontre par récurrence sur  $i$ . Pour  $i = 1$ , on a  $y_1 = y, z_1 = z$  et le résultat est évident. Supposons le résultat vrai pour  $j < i$ , alors

$$|y_i - z_i| \leq |y - z| + h \sum_{j < i} |a_{ij}| k \max_{j < i} |y_j - z_j|,$$

$$|y_i - z_i| \leq |y - z| + \alpha k h \max_{j < i} |y_j - z_j|.$$

Par hypothèse de récurrence il vient

$$\max_{j < i} |y_j - z_j| \leq (1 + \alpha k h + \dots + (\alpha k h)^{i-2}) |y - z|,$$

et l'inégalité s'ensuit à l'ordre  $i$ . ■

La formule (2.8) entraîne maintenant

$$|\Phi(t, y, h) - \Phi(t, z, h)| \leq \sum_{1 \leq j \leq q} |b_j| k |y_j - z_j| \leq \Lambda |y - z|,$$

avec

$$\Lambda = k \sum_{1 \leq j \leq q} |b_j| (1 + (\alpha k h_{\max}) + \dots + (\alpha k h_{\max})^{j-1}),$$

d'où le corollaire suivant

**Corollaire 2.19** *Les méthodes de Runge-Kutta sont stables, avec constante de stabilité  $S = e^{\Lambda t}$ .*

#### 2.4.4 Ordre des méthodes de Runge-Kutta

Pour déterminer l'ordre, on applique la Propositions 2.10 : l'ordre est au moins égal à  $p$  si et seulement si

$$\frac{\partial^\ell \Phi}{\partial h^\ell}(t, y, 0) = \frac{1}{\ell + 1} f^{[\ell]}(t, y), \quad \ell \leq p - 1.$$

Grâce à l'expression (2.8), on obtient facilement les dérivées successives de  $\Phi$  :

$$\Phi(t, y, 0) = \sum_{1 \leq j \leq q} b_j f(t, y) = f(t, y).$$

D'après le Théorème 2.6, les méthodes de Runge-Kutta sont donc toujours d'ordre  $\geq 1$ , i.e. consistantes.

**Lemme 2.20**

$$\frac{\partial \Phi}{\partial h}(t, y, 0) = \left( \sum_{i=1}^q b_i c_i \right) f^{[1]}(t, y)$$

**Preuve :** On a

$$\Phi(t, y, h) = \sum_{i=1}^q b_i f(t + c_i h, y_i(t, y, h)),$$

d'où

$$\frac{\partial \Phi}{\partial h}(t, y, h) = \sum_{i=1}^q b_i c_i \frac{\partial f}{\partial t}(t + c_i h, y_i(t, y, h)) + \sum_{i=1}^q b_i \frac{\partial y_i}{\partial h} \frac{\partial f}{\partial y}(t + c_i h, y_i(t, y, h)),$$

Par ailleurs

$$y_i(t, y, h) = y + h \sum_{j=1}^{i-1} a_{ij} f(t + c_j h, y_j(t, y, h)),$$

donc

$$\frac{\partial y_i}{\partial h}(t, y, h) = \sum_{j=1}^{i-1} a_{ij} f(t + c_j h, y_j(t, y, h)) + h \sum_{j=1}^{i-1} a_{ij} \frac{\partial}{\partial h} \left( f(t + c_j h, y_j(t, y, h)) \right).$$

Puisque  $y_j(t, y, 0) = y$  et grâce à (2.7) on obtient

$$\frac{\partial y_i}{\partial h}(t, y, 0) = \left( \sum_{j=1}^{i-1} a_{ij} \right) f(t, y) = c_i f(t, y),$$

par conséquent

$$\begin{aligned} \frac{\partial \Phi}{\partial h}(t, y, 0) &= \left( \sum_{i=1}^q b_i c_i \right) \left( \frac{\partial f}{\partial t}(t, y) + f(t, y) \frac{\partial f}{\partial y}(t, y) \right) \\ &= \left( \sum_{i=1}^q b_i c_i \right) f^{[1]}(t, y). \end{aligned}$$

■

**Corollaire 2.21** Les méthodes de Runge-Kutta sont d'ordre  $\geq 2$  si et seulement si  $\sum_{i=1}^q b_i c_i = \frac{1}{2}$ .

**Exemple 2.22**

- On reprend le tableau de la méthode (RK2) : la méthode est d'ordre au moins 2 puisque  $0 \times (1 - \frac{1}{2\alpha}) + \alpha \times \frac{1}{2\alpha} = \frac{1}{2}$ .

- La méthode classique (RK4) est d'ordre au moins 2 également, puisque  $0 \times \frac{1}{6} + \frac{1}{2} \times \frac{2}{6} \times 1 \times \frac{1}{6} = \frac{1}{2}$ .

Un calcul (assez pénible!) de  $\frac{\partial^2 \Phi}{\partial h^2}$  et de  $\frac{\partial^3 \Phi}{\partial h^3}$  conduirait au résultat suivant :

**Théorème 2.23** Les méthodes de Runge-Kutta définies par les tableaux de coefficients  $c_i, a_{ij}, b_j$  sont

- d'ordre  $\geq 2$  ssi  $\sum_{i=1}^q b_i c_i = \frac{1}{2}$ .
- d'ordre  $\geq 3$  ssi  $\sum_{i=1}^q b_i c_i = \frac{1}{2}$ ;  $\sum_{i=1}^q b_i c_i^2 = \frac{1}{3}$ ;  $\sum_{i,j} b_i a_{ij} c_j = \frac{1}{6}$ .

- $d'ordre \geq 4$  ssi

$$\begin{aligned} \sum_{i=1}^q b_i c_i &= \frac{1}{2}; \quad \sum_{i=1}^q b_i c_i^2 = \frac{1}{3}; \quad \sum_{i=1}^q b_i c_i^3 = \frac{1}{3}; \\ \sum_{i,j} b_i a_{ij} c_j &= \frac{1}{6}; \quad \sum_{i,j} b_i a_{ij} c_j^2 = \frac{1}{12} \quad \sum_{i,j} b_i c_i a_{ij} c_j = \frac{1}{8}; \\ \sum_{i,j} b_i a_{ij} a_{jk} c_k &= \frac{1}{12}. \end{aligned}$$

**Exemple 2.24** Par un calcul simple du tableau, on peut vérifier que la méthode classique (RK4) est d'ordre au moins 4.

### Remarques 2.25

- i) Les méthodes de Runge-Kutta ne nécessitent pas le calcul des dérivées successives de  $f$ , contrairement aux méthodes de Taylor. Elles donnent de très bon résultats numériques et sont les plus utilisées pour résoudre les problèmes de Cauchy.
- ii) Il est possible d'obtenir des méthodes d'ordre inférieur au rang  $q$ . On peut démontrer ([1]) que pour avoir une méthode de Runge-Kutta d'ordre 5, il faut une méthode de rang  $q = 6$ , ce qui fait du schéma (RK4) l'algorithme de Runge-Kutta le plus utilisé.

## 2.5 Application aux systèmes d'équations différentielles

La forme générale d'un système de  $m$  équations différentielles avec conditions initiales s'écrit :

$$\begin{cases} y_1'(t) = f_1(t, y_1(t), y_1(t), \dots, y_m(t)) & (y_1(t_0) = y_{1,0}) \\ y_2'(t) = f_2(t, y_1(t), y_1(t), \dots, y_m(t)) & (y_2(t_0) = y_{2,0}) \\ \vdots & \vdots \\ y_m'(t) = f_m(t, y_1(t), y_1(t), \dots, y_m(t)) & (y_m(t_0) = y_{m,0}) \end{cases}$$

Ici encore, on note  $y_i(t_n)$ , la valeur exacte de la  $i^e$  variable dépendante en  $t = t_n$  et  $y_{i,n}$ , son approximation numérique.

Parmi les techniques de résolution de systèmes d'équations différentielles, nous ne présentons que la méthode de Runge-Kutta d'ordre 4. Il est possible d'utiliser également les autres méthodes déjà vues, ou que nous allons voir.

### 2.5.1 Algorithme (RK4)

1. Etant donné un pas de temps  $h$  et un nombre maximal d'itérations  $N$ .
2. Etant données des conditions initiales  $t_0, y_{1,0}, \dots, y_{m,0}$ .

3. Pour  $0 \leq n < N$  :

$$\text{pour } 1 \leq i \leq m, k_{i,n,1} = hf_i(t_n, y_{1,n}, \dots, y_{m,n})$$

$$\text{pour } 1 \leq i \leq m, k_{i,n,2} = hf_i\left(t_n + \frac{h}{2}, y_{1,n} + \frac{k_{1,n,1}}{2}, \dots, y_{m,n} + \frac{k_{m,n,1}}{2}\right)$$

$$\text{pour } 1 \leq i \leq m, k_{i,n,3} = hf_i\left(t_n + \frac{h}{2}, y_{1,n} + \frac{k_{1,n,2}}{2}, \dots, y_{m,n} + \frac{k_{m,n,2}}{2}\right)$$

$$\text{pour } 1 \leq i \leq m, k_{i,n,4} = hf_i\left(t_n + h, y_{1,n} + k_{1,n,3}, \dots, y_{m,n} + k_{m,n,3}\right)$$

$$\text{pour } 1 \leq i \leq m, y_{i,n+1} = y_{i,n} + \frac{1}{6}(k_{i,n,1} + 2k_{i,n,2} + 2k_{i,n,3} + k_{i,n,4})$$

$$t_{n+1} = t_n + h.$$

Cet algorithme est complexe en apparence. Pour bien en comprendre le principe, il suffit d'y voir une application de la méthode de Runge-Kutta à chacune des équations différentielles. De plus, il est nécessaire de calculer les  $m$  constantes  $k_{i,n,1}$  avant de passer au calcul des constantes  $k_{i,n,2}$  et ainsi de suite.

### 2.5.2 Exemple

Soit le système de deux équations différentielles suivant :

$$\begin{cases} y_1'(t) = y_2(t) & (y_1(0) = 2) \\ y_2'(t) = 2y_2(t) - y_1(t) & (y_2(0) = 1) \end{cases}$$

Dont la solution analytique est :

$$y_1(t) = 2e^t - te^t$$

$$y_2(t) = e^t - te^t$$

On a alors :

$$f_1(t, y_1(t), y_2(t)) = y_2(t)$$

$$f_2(t, y_1(t), y_2(t)) = 2y_2(t) - y_1(t)$$

et la condition initiale  $(t_0, y_{1,0}, y_{2,0}) = (0, 2, 1)$ . Si l'on prend par exemple  $h = 0.1$ , on trouve :



$$\begin{aligned}
k_{1,0,1} &= 0.1(f_1(0, 2, 1)) = 0.1 \\
k_{2,0,1} &= 0.1(f_2(0, 2, 1)) = 0 \\
k_{1,0,2} &= 0.1(f_1(0.05, 2.05, 1)) = 0.1 \\
k_{2,0,2} &= 0.1(f_2(0.05, 2.05, 1)) = -0.005 \\
k_{1,0,3} &= 0.1(f_1(0.05, 2.05, 0.9975)) = 0.09975 \\
k_{2,0,3} &= 0.1(f_2(0.05, 2.05, 0.9975)) = -0.0055 \\
k_{1,0,4} &= 0.1(f_1(0.1, 2.09975, 0.9945)) = 0.09945 \\
k_{2,0,4} &= 0.1(f_2(0.1, 2.09975, 0.9945)) = -0.011075 \\
y_{1,1} &= y_{1,0} + \frac{1}{6}(0.1 + 2(0.1) + 2(0.09975) + 0.09945) \\
&= 2.099825 \\
y_{2,1} &= y_{2,0} + \frac{1}{6}(0 + 2(-0.005) + 2(-0.0055) + (-0.011075)) \\
&= 0.994651667.
\end{aligned}$$

### 3 Quelques exemples d'application

**Exemple 3.1** (*Loi de Newton*) : Considérons une tasse de café à la température de  $75^\circ\text{C}$  dans une salle à  $25^\circ\text{C}$ . On suppose que la température du café suit la loi de Newton, c'est-à-dire que la vitesse de refroidissement du café est proportionnelle à la différence des températures. En formule, cela signifie qu'il existe une constante  $K < 0$  telle que la température vérifie l'équation différentielle ordinaire du premier ordre

$$T'(t) = K(T(t) - 25).$$

La condition initiale ( $I$ ) est donc simplement

$$T(0) = 75.$$

Pour calculer la température à chaque instant on a besoin de connaître la constante  $K$ . Cette valeur peut être déduite en constatant qu'après 5 minutes, le café est à  $50^\circ\text{C}$ , c'est-à-dire

$$T(5) = 50.$$

Calculer la solution exacte de ce problème de Cauchy et la comparer avec la solution approchée obtenue par la méthode d'Euler explicite.

**Solution :** On commence par calculer toutes les solutions de l'EDO. Étant une équation différentielle du premier ordre, la famille de solutions dépendra d'une constante qu'on fixera en utilisant la condition initiale ( $I$ ). Il s'agit d'une EDO à variables séparables donc formellement on a

$$T'(t) = K(T(t) - 25)$$

implique que

$$\frac{dT}{T(t) - 25} = K dt$$

d'où

$$T(t) = 25 + 50e^{Kt}.$$

Il ne reste qu'à établir la valeur numérique de la constante de refroidissement  $K$  grâce à l'«indice»

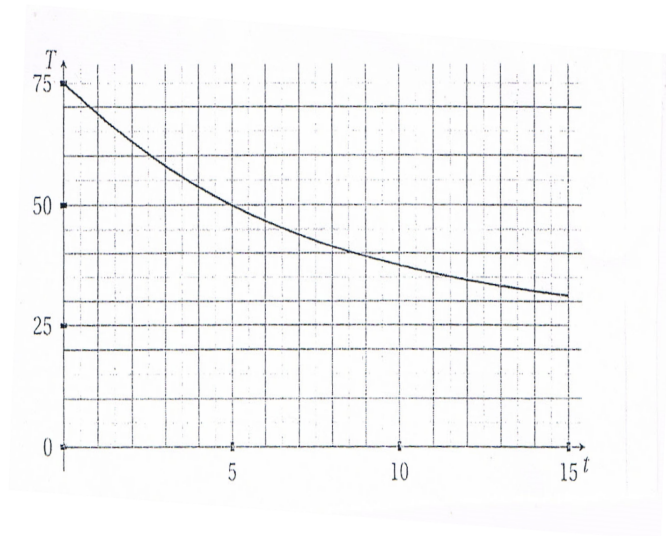
$$50 = T(5) = 25 + 50e^{5K},$$

d'où

$$K = -\frac{\ln(2)}{5} \simeq -0.14,$$

c'est-à dire que la température du café évolue selon la fonction

$$T(t) = 25 + 50e^{-\frac{\ln(2)}{5}t}.$$



**Solution approchée par le schéma d'Euler explicite :** Supposons de connaître  $K$  mais de ne pas vouloir/pouvoir calculer la fonction  $T$ . Grâce à la méthode d'Euler on peut estimer la température à différentes instantes  $t_n$  en faisant une discrétisation temporelle du futur, i.e. on construit une suite de valeurs  $(t_n = nh_n)_n$  et en construisant une suite de valeurs  $(T_n)_n$  où chaque  $T_n$  est une approximation de  $T(t_n)$ . Si on utilise la méthode d'Euler explicite avec un pas de temps fixé  $h$ , cette suite de température est ainsi construite :

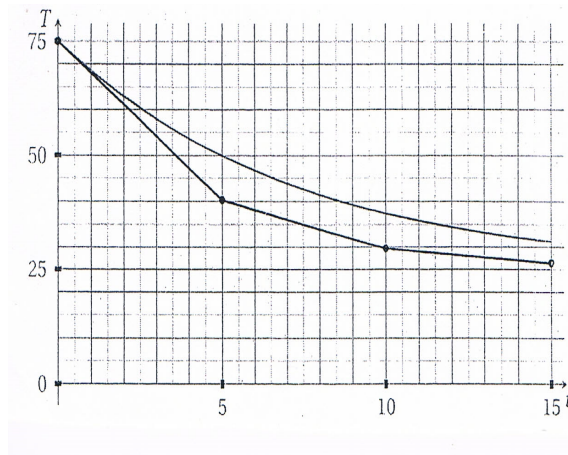
$$\begin{cases} T_{n+1} = T_n - \frac{\ln(2)}{5}h(T_n - 25), \\ T_0 = 75, \end{cases}$$

qu'on peut réécrire comme

$$\begin{cases} T_{n+1} = (1 - \frac{\ln(2)}{5}h)T_n + 5\ln(2)h, \\ T_0 = 75. \end{cases}$$

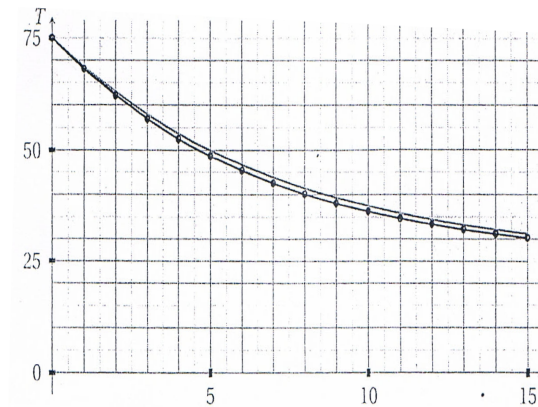
- Si on prend  $h = 5$ , on obtient le tableau suivant :

$t_n$	$T(t_n)$	$T_n$	$T(t_n) - T_n$
0.000000	75.000000	75.000000	0.000000
5.000000	50.000000	40.342641	9.657359
10.000000	37.500000	29.707933	7.792067
15.000000	31.250000	26.444642	4.805358



- Si on prend  $h = 1$ , on obtient le tableau suivant :

$t_n$	$T(t_n)$	$T_n$	$T(t_n) - T_n$
0.000000	75.000000	75.000000	0.000000
1.000000	68.527528	68.068528	0.459000
2.000000	62.892914	62.097962	0.794952
3.000000	57.987698	56.955093	1.032605
4.000000	53.717459	52.525176	1.192283
5.000000	50.000000	48.709377	1.290623
6.000000	46.763764	45.422559	1.341205
7.000000	43.946457	42.591391	1.355066
8.000000	41.493849	40.152707	1.341142
9.000000	39.358729	38.052095	1.306634
10.000000	37.500000	36.242691	1.257309
11.000000	35.881882	34.684123	1.197759
12.000000	34.473229	33.341618	1.131610
13.000000	33.246924	32.185225	1.061700
14.000000	32.179365	31.189141	0.990224
15.000000	31.250000	30.331144	0.918856



### Exemple 3.2

Soit le problème de Cauchy suivant

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

On rappelle que l'algorithme de Taylor d'ordre  $p$  à pas de temps constant  $h$  s'écrit :

$$y_{n+1} = y_n + hT_p(t_n, y_n, h),$$

avec  $T_p(t, y, h) = \sum_{k=1}^p \frac{h^{k-1}}{k!} f^{[k-1]}(t, y).$

On propose l'algorithme suivant, appelé "mixed Runge-Kutta-Taylor method" :

$$y_{n+1} = y_n + hG_3(t_n, y_n, h),$$

avec

$$G_3(t_n, y_n, h) = f(t_n, y_n) + \frac{h}{2}f^{[1]}(t_n + \frac{h}{3}, y_n + \frac{h}{3}f(t_n, y_n)).$$

1) En comparant  $G_3$  et  $T_3$ , montrer que  $G_3$  est aussi d'ordre 3. Quel avantage selon vous présente  $G_3$  par rapport à la méthode de Taylor d'ordre 3.?

2) On se propose de tester cette méthode sur l'exemple très simple suivant

$$y'(t) = 1 - y(t), \text{ avec } y_0 = 2 \text{ et } t_n = nh.$$

a) Calculer explicitement  $y_n$  dans la méthode  $G_3$  en fonction de  $n$  et  $h$ .

b) Prenons  $h = 0.1$ , calculer  $y_5$  et la comparer à la valeur  $y(0.5)$  de la solution exacte.

**Solution :**

1) Appliquons la formule de Taylor à  $f^{[1]}$  :

$$f^{[1]}(t + \frac{h}{3}, y + \frac{h}{3}f(t, y)) = f^{[1]}(t, y) + \frac{h}{3} \frac{\partial f^{[1]}}{\partial t}(t, y) + \frac{h}{3} f(t, y) \frac{\partial f^{[1]}}{\partial y}(t, y) + o(h^2)$$

On reconnaît dans le membre de droite l'expression de  $f^{[2]} = \frac{\partial f^{[1]}}{\partial t} + f \frac{\partial f^{[1]}}{\partial y}$ , qui permet d'écrire

$$\begin{aligned}
G_3(t, y, h) &= f(t, y) + \frac{h}{2} \left( f^{[1]}(t, y) + \frac{h}{3} \frac{\partial f^{[1]}}{\partial t}(t, y) + \frac{h}{3} f(t, y) \frac{\partial f^{[1]}}{\partial y}(t, y) \right) + o(h^3) \\
&= f(t, y) + \frac{h}{2} f^{[1]}(t, y) + \frac{h^2}{6} f^{[2]}(t, y) + o(h^3) = T_3(t, y, h) + o(h^3).
\end{aligned}$$

On en déduit que  $G_3$  est d'ordre 3 comme le schéma de Taylor ( $\mathcal{T}_3$ ) (voir Exemple 2.9). L'avantage de  $G_3$  est que la fonction  $\Phi$  peut s'exprimer sans avoir recours à la fonction  $f^{[2]}$ , c'est à dire aux dérivées secondes de  $f$ .

2)

a) On a  $f(t, y) = 1 - y$  et  $f^{[1]}(t, y) = \frac{\partial f}{\partial t} + (1 - y) \frac{\partial f}{\partial y} = y - 1$ , et il vient donc pour  $G_3$  :

$$G_3(t, y, h) = (1 - y) + \frac{h}{2}(y + \frac{h}{3}(y - 1) - 1) = 1 - y + \frac{h}{2}(y - 1) + \frac{h^2}{6}(1 - y)$$

On a donc

$$\begin{aligned}
y_{n+1} &= y_n + h - hy_n + \frac{h^2}{2}(y_n - 1) + \frac{h^3}{6}(1 - y_n) \\
&= (1 - h + \frac{h^2}{2} - \frac{h^3}{6})y_n + h - \frac{h^2}{2} + \frac{h^3}{6}.
\end{aligned}$$

Ceci revient à

$$y_{n+1} - 1 = (1 - \alpha)y_n + \alpha - 1 = (1 - \alpha)(y_n - 1)$$

où  $\alpha = h - \frac{h^2}{2} + \frac{h^3}{6}$ . Donc par récurrence on obtient

$$y_n = 1 + (1 - \alpha)^n.$$

b) **Application numérique** : pour  $h = 0.1$ , calculer la solution approchée de la solution exacte en  $t = 0.5$  revient à calculer  $y_5$  :

la solution exacte du problème est  $y(t) = 1 + e^{-t}$ ; et on trouve  $y_5 = 1.606517$ , contre  $y(0.5) = 1.606530$ .

### Exemple 3.3

En donnant les solutions de l'équation différentielle ci-dessous avec la condition initiale  $y(0) = 1$ , puis  $y(0) = 1 + \varepsilon$ ,  $\varepsilon$  un réel non nul. vérifier qu'elle conduit à des schémas instables.

$$\begin{cases} y'(t) = 36y(t) - 37e^{-t} \\ y(0) = 1, \text{ puis } 1 + \varepsilon \end{cases}$$

**Solution** : En appliquant la méthode de la variation de la constante, on trouve la solution générale

$$y(t) = (e^{-37} + C)e^{36t},$$

où  $C$  est la constante d'intégration :

- Si  $y(0) = 1$ , alors  $C = 0$ , et la solution du problème est  $y(t) = e^{-t}$ .
- Si  $y(0) = 1 + \varepsilon$ , alors  $C = \varepsilon$  et la solution du problème est  $y_\varepsilon(t) = e^{-t} + \varepsilon e^{36t}$ .

**Conclusion** : En comparant  $y(t)$  et  $y_\varepsilon(t)$ , on voit que la différence  $y(t) - y_\varepsilon(t) = \varepsilon e^{36t}$ . Même si  $\varepsilon$  est très petit, cet écart tend vers  $+\infty$  avec  $t$ . Les deux solutions divergent l'une de l'autre. Ce problème est donc très sensible aux conditions initiales.

### Exemple 3.4

1) Montrer que le problème de Cauchy

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

possède une unique solution, donner son expression explicite.

2) Calculer une valeur approchée de  $y(1)$  en utilisant le schéma de Runge-Kutta d'ordre 4 avec un pas  $h$  égal à 0.1.

#### Solution :

1) Il est clair que  $f(t, y) = 1 + y$  est continue sur  $[0, 1] \times \mathbb{R}$  et que  $|f(t, y_1) - f(t, y_2)| = |y_1 - y_2|$ , c-à-d  $f$  est lipschitzienne avec une constante de Lipschitz égal à 1. Les hypothèses du théorème de Cauchy-Lipschitz sont vérifiées, le problème de Cauchy a une solution unique. On montre immédiatement que cette solution est

$$y(t) = e^t - 1.$$

2) La méthode (RK4) donne sur cet exemple

$$\begin{aligned} p_{n,1} &= f(t_n, y_n) = 1 + y_n \\ p_{n,2} &= f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}p_{n,1}\right) = 1 + y_n + \frac{h}{2}p_{n,1} \\ p_{n,3} &= f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}p_{n,2}\right) = 1 + y_n + \frac{h}{2}p_{n,2} \\ p_{n,4} &= f(t_n + h, y_n + hp_{n,3}) = 1 + y_n + hp_{n,3} \\ y_{n+1} &= y_n + \frac{h}{6}(p_{n,1} + 2p_{n,2} + 2p_{n,3} + p_{n,4}). \end{aligned}$$

Rassemblons dans un tableau les valeurs approchées obtenues, les valeurs exactes et l'erreur effective  $e_n = y(t_n) - y_n$  :

$t_n$	$y_n$	$y(t_n)$	$e_n$
0.000000	0.000000	0.000000	0.000000
0.1	0.10517083	0.10517092	$8.5 \times 10^{-8}$
0.2	0.22140257	0.22140276	$1.9 \times 10^{-7}$
0.3	0.34985850	0.34985881	$3.1 \times 10^{-7}$
0.4	0.49182424	0.49182470	$4.6 \times 10^{-7}$
0.5	0.64872064	0.64872127	$6.3 \times 10^{-7}$
0.6	0.82211796	0.82211880	$8.4 \times 10^{-7}$
0.7	1.01375163	1.01375271	$1.1 \times 10^{-6}$
0.8	1.22553956	1.22554093	$1.4 \times 10^{-6}$
0.9	1.45960141	1.45960311	$1.7 \times 10^{-6}$
1	1.71827974	1.71828183	$2.1 \times 10^{-6}$

### Exemple 3.5

Soit l'équation différentielle du second ordre à conditions initiales :

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

1) Ecrire cette équation différentielle sous la forme d'un système différentiel de deux équations différentielles d'ordre un.

2) Appliquer la méthode d'Euler puis la méthode de Runge-Kutta (RK2) à ce système.

**Solution :**

1) On pose  $y' = z$ , d'où  $y'' = z' = -2z + 2y$ . On trouve le système

$$\begin{cases} y' = z \\ z' = 2y - 2z \\ y(0) = 1, y'(0) = 2 \end{cases}$$

Matriciellement ce système s'écrit

$$\begin{cases} U' = AU = F(t, U) \\ U(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{cases}$$

avec  $U = \begin{pmatrix} y \\ z \end{pmatrix}$ ,  $A = \begin{pmatrix} 0 & 1 \\ 2 & -2 \end{pmatrix}$  et  $F$  est l'endomorphisme associé à la matrice  $A$ .

2)

a) Appliquons la méthode d'Euler à ce système :

$$U_{n+1} = U_n + h_n F(t_n, U_n)$$

D'où

$$\begin{aligned} U_{n+1} &= U_n + h_n A U_n \\ &= \begin{pmatrix} y_n \\ z_n \end{pmatrix} + h_n \begin{pmatrix} 0 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} y_n \\ z_n \end{pmatrix}. \end{aligned}$$

On obtient donc l'algorithme

$$\begin{cases} y_{n+1} = y_n + h_n z_n \\ z_{n+1} = z_n + 2h_n(y_n - z_n) \end{cases}$$

avec  $y_0 = 1$  et  $z_0 = 2$ .

b) Appliquons maintenant la méthode (RK2) :

$$\begin{aligned} U_{n+1} &= U_n + h_n F(t_n + \frac{h_n}{2}, U_n + \frac{h_n}{2} F(t_n, U_n)) \\ &= \begin{pmatrix} y_n \\ z_n \end{pmatrix} + h_n \begin{pmatrix} 0 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} y_n + \frac{h_n}{2} z_n \\ z_n + h_n(y_n - z_n) \end{pmatrix} \\ &= \begin{pmatrix} y_n + z_n + h_n(y_n - z_n) \\ -z_n + 2y_n + h_n(3z_n - 2y_n) \end{pmatrix} \end{aligned}$$

On obtient donc l'algorithme

$$\begin{cases} y_{n+1} = y_n + h_n(y_n - z_n) \\ z_{n+1} = -z_n + 2y_n + h_n(3z_n - 2y_n) \end{cases}$$

avec  $y_0 = 1$  et  $z_0 = 2$ .

## 4 Schémas numériques multi-pas

On s'intéresse toujours à la résolution numérique du problème de Cauchy  $\{(E), (I)\}$ , où  $f$  est une fonction continue de  $[t_0, t_0 + T] \times \mathbb{R}$  dans  $\mathbb{R}$ ,  $k$ -lipschitzienne en  $y$ . On considère une subdivision  $(t_n)_{0 \leq n \leq N}$  de l'intervalle  $[t_0, t_0 + T]$  de pas successifs  $h_n = t_{n+1} - t_n$ .

**Définition 4.1** *Un schéma numérique à  $r + 1$  pas est de la forme*

$$\begin{cases} y_{n+1} = \Phi(y_{n+1}, t_n, y_n, h_n \dots, t_{n-r}, y_{n-r}, h_{n-r}), & r \leq n \leq N - 1 \\ t_{n+1} = t_n + h_n \end{cases} \quad (4.1)$$

où  $r$  est un entier  $\geq 0$ .

**Exemple 4.2** *Une classe de méthodes à  $r + 1$  pas dites linéaires à pas de discrétisation  $h_n = h$  constant est définie par le schéma*

$$y_{n+1} = \sum_{i=0}^r \alpha_i y_{n-i} + h \sum_{i=-1}^r \beta_i f_{n-i}, \quad r \leq n \leq N - 1, \quad (4.2)$$

dans lequel les valeurs  $f_{n-i} = f(t_{n-i}, y_{n-i})$ , interviennent de manière linéaire, ce qui n'était pas le cas pour les méthodes de Runge-Kutta. La relation de récurrence (4.2) peut s'écrire sous une autre forme dite d'équation aux différences linéaire

$$\sum_{i=0}^{r+1} a_i y_{n+i} = h \sum_{i=0}^{r+1} b_i f_{n+i}, \quad 0 \leq n \leq N - r - 1, \quad (4.3)$$

en posant  $\alpha_i = -\frac{a_{r-i}}{a_{r+1}}$  et  $\beta_i = \frac{b_{r-i}}{a_{r+1}}$ , où les constantes  $a_i, b_i$  sont indépendantes de  $n$ ,  $a_{r+1} \neq 0$ .

- Si on considère la forme (4.2), le point initial  $(t_0, y_0)$  étant donné ; l'algorithme ne peut démarrer que si les valeurs  $(y_1, f_1), \dots, (y_r, f_r)$  ont déjà été calculées. Ce calcul ne peut être fait que par une méthode à un pas pour  $(y_1, f_1)$ , à au plus 2 pas pour  $(y_2, f_2), \dots$  au plus  $r$  pas pour  $(y_r, f_r)$ . On dit que les méthodes à pas multiples **ne sont pas auto-démarrantes**. L'initialisation des  $r$  premières valeurs  $(y_i, f_i), 1 \leq i \leq r$ , sera généralement faite à l'aide d'une méthode de Runge-Kutta d'ordre supérieur ou égal à celui de la méthode (4.2).
- Si  $\beta_{-1} = 0$ , nous pouvons obtenir directement la valeur de  $y_{n+1}$  en fonction de  $y_n, \dots, y_{n-r}$ ; la méthode est **explicite**.
- Si  $\beta_{-1} \neq 0$ , la méthode est **implicite** car elle définit  $y_{n+1}$  par une équation implicite. Nous allons voir une méthode itérative permettant d'approcher la solution de cette équation. Plus précisément nous avons le

**Théorème 4.3** *Si  $\beta_{-1} \neq 0$ , l'équation (4.2) admet une solution unique si*

$$h < \frac{1}{k\beta_{-1}},$$

où  $f$  est supposé toujours  $k$ -lipschitzienne en  $y$ . La suite  $(y_n^{(p)})_p$  définie par la méthode itérative

$$y_{n+1}^{(p+1)} = h\beta_{-1}f(t_{n+1}, y_{n+1}^{(p)}) + \sum_{i=0}^r (\alpha_i y_{n-i} + h\beta_i f_{n-i}) \quad (4.4)$$



converge vers  $y_{n+1}$  lorsque  $p$  tend vers l'infini.

**Preuve :** Si nous supposons connues les valeurs  $y_{n-r}, \dots, y_n$  et les valeurs  $f_{n-r}, \dots, f_n$ , le deuxième terme du second membre de (4.4) est une constante que nous noterons  $C$ . Donc  $y_{n+1}$  est solution de l'équation

$$y = h\beta_{-1}f(t_{n+1}, y) + C$$

En utilisant le théorème du point fixe (voir annexe) et la méthode des approximations successives pour la fonction  $\varphi : y \mapsto h\beta_{-1}f(t_{n+1}, y) + C$  supposée contractante

$$\begin{aligned} |\varphi(y) - \varphi(y^*)| &\leq h|\beta_{-1}||f(t_{n+1}, y) - f(t_{n+1}, y^*)| \\ &\leq hk|\beta_{-1}||y - y^*| = K|y - y^*|, \quad K < 1 \end{aligned}$$

nous obtenons la condition  $K = hk|\beta_{-1}| < 1$ . En outre, on a l'estimation de l'erreur

$$|y_{n+1}^{(p)} - y_{n+1}| \leq (hk|\beta_{-1}|)^p \frac{1}{1 - hk|\beta_{-1}|} |y_{n+1}^{(1)} - y_{n+1}|.$$

■

## 4.1 Consistance, ordre, stabilité et convergence

**Définition 4.4** Soit  $y$  une solution exacte de l'équation (E). L'erreur locale de consistance  $e_n$  relative à  $y$  est définie par

$$e_n = y(t_{n+1}) - \bar{y}_{n+1}, \quad r \leq n \leq N-1,$$

où  $\bar{y}_{n+1}$  est l'approximation fournie par la méthode (4.1) sous l'hypothèse  $y_{n-i} = y(t_{n-i})$ ,  $i = 0, \dots, r$ .

On dit que la méthode est consistante si pour toute solution exacte  $y$  la somme des erreurs de consistance relatives à  $y$ , soit  $\sum_{n=r}^{N-1} |e_n|$ , tend vers 0 quand  $h_{\max}$  tend vers 0.

La méthode est dite d'ordre  $p$  si pour toute solution  $y$  il existe une constante  $C$  telle que

$$|e_n| \leq Ch_n h_{\max}^p.$$

### Exemple 4.5

Déterminons l'erreur de consistance  $e_n$  dans le cas de la méthode (4.2) quand  $\beta_{-1} = 0$ . Supposons que  $f$  est de classe  $C^p$ ,  $y$  est alors de classe  $C^{p+1}$ . Par ailleurs

$$\begin{aligned} y(t_{n+1}) &= y(t_n + h) = \sum_{k=0}^p \frac{h^k}{k!} y^{(k)}(t_n) + O(h^{p+1}) \\ y(t_{n-i}) &= y(t_n - ih) = \sum_{k=0}^p \frac{(-ih)^k}{k!} y^{(k)}(t_n) + O(h^{p+1}), \\ f(t_{n-i}, y(t_{n-i})) &= y'(t_n - ih) = \sum_{k=0}^{p-1} \frac{(-ih)^k}{k!} y^{(k+1)}(t_n) + O(h^p) \\ &= \sum_{k=0}^p \frac{(-ih)^{k-1}}{k!} y^{(k)}(t_n) + O(h^p). \end{aligned}$$

Par conséquent

$$\begin{aligned}
e_n &= y(t_{n+1}) - \sum_{i=0}^r \left( \alpha_i y(t_{n-i}) + h \beta_i f(t_{n-i}, y(t_{n-i})) \right) \\
&= \sum_{k=0}^p \frac{h^k}{k!} y^{(k)}(t_n) \left( 1 - \sum_{i=0}^r (\alpha_i (-i)^k + k \beta_i (-i)^{k-1}) \right) + O(h^{p+1}) \\
&= \sum_{k=0}^p \frac{h^k}{k!} y^{(k)}(t_n) \left( 1 - (-1)^k \sum_{i=0}^r (i^k \alpha_i - k i^{k-1} \beta_i) \right) + O(h^{p+1}).
\end{aligned}$$

On constate donc que la méthode est d'ordre  $p$  si et seulement si

$$\sum_{i=0}^r (i^k \alpha_i - k i^{k-1} \beta_i) = (-1)^k, \quad 0 \leq k \leq p.$$

En particulier, elle est consistante si et seulement si

$$\begin{cases} \alpha_0 + \alpha_1 + \dots + \alpha_r = 1 \\ \alpha_1 + \dots + r \alpha_r - (\beta_0 + \dots + \beta_r) = -1. \end{cases}$$

**Définition 4.6** Soient  $(y_n)$  et  $(\tilde{y}_n)$  deux suites définies par les équations

$$y_{n+1} = \Phi(t_{n-i}, y_{n-i}, h_{n-i}), \quad r \leq n \leq N-1, \quad (4.5)$$

$$\tilde{y}_{n+1} = \Phi(t_{n-i}, \tilde{y}_{n-i}, h_{n-i}) + \varepsilon_n, \quad r \leq n \leq N-1, \varepsilon_n \in \mathbb{R}. \quad (4.6)$$

On dit que le schéma (4.6) est le schéma (4.5) perturbé par  $\varepsilon_n$  et par les nouvelles conditions initiales  $\tilde{y}_0, \dots, \tilde{y}_r$ .

**Définition 4.7** La méthode à  $r+1$  pas (4.5) est stable s'il existe une constante dite de stabilité  $S$  telle que

$$\max_{0 \leq n \leq N} |y_n - \tilde{y}_n| \leq S \left( \max_{0 \leq n \leq r} |y_n - \tilde{y}_n| + \sum_{n=r}^{N-1} |\varepsilon_n| \right).$$

**Définition 4.8** La méthode à  $r+1$  pas (4.1) est convergente si

$$\left( \max_{0 \leq n \leq r} |y_n - y(t_n)| \rightarrow 0 \text{ quand } h_{\max} \rightarrow 0 \right) \Rightarrow \left( \max_{0 \leq n \leq N} |y_n - y(t_n)| \rightarrow 0 \text{ quand } h_{\max} \rightarrow 0 \right).$$

En appliquant la définition (4.6) à  $\tilde{y}_n = y(t_n)$ , on voit que l'erreur globale de la suite  $(y_n)$  admet la majoration

$$\max_{0 \leq n \leq N} |y_n - y(t_n)| \leq S \left( \max_{0 \leq n \leq r} |y_n - y(t_n)| + \sum_{n=r}^{N-1} |e_n| \right). \quad (4.7)$$

Si on suppose que la méthode est consistante, on constate donc qu'elle est convergente, d'où le théorème suivant

**Théorème 4.9** Si la méthode à  $r+1$  pas (4.1) est consistante et stable, elle est convergente.

**Remarque 4.10** Si la méthode (4.1) est d'ordre  $p$ , i.e.  $|e_n| \leq C h_n h_{\max}^p$ , alors on a  $\sum_{n=r}^{N-1} |e_n| \leq$

$C T h_{\max}^p$ , car  $\sum_{n=0}^{N-1} h_n = T$ . Pour la phase d'initialisation, il convient donc de choisir une méthode dont l'erreur globale d'initialisation  $\max_{0 \leq n \leq r} |y_n - y(t_n)|$  est de l'ordre de  $h_{\max}^p$ . Ceci conduit à choisir une méthode d'initialisation d'ordre  $\geq p-1$ .

## 4.2 Schémas d'Adams-Baschforth

Si  $y$  est une solution exacte de l'équation différentielle  $y'(t) = f(t, y(t))$  alors

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt.$$

Les schémas d'Adams approchent l'intégrale  $\int_{t_n}^{t_{n+1}} f(t, y(t)) dt$  par l'intégrale d'un polynôme  $P$  interpolant  $f$  en des points donnés qui peuvent être à l'extérieur de l'intervalle  $[t_n; t_{n+1}]$ . On peut construire différents schémas selon les points d'interpolation choisis. Ils se divisent en deux familles : les méthodes d'Adams-Bashforth qui sont explicites et les méthodes d'Adams-Moulton qui sont implicites.

**-Méthodes d'Adams-Baschforth :** Il s'agit de méthodes explicites à  $r + 1$  pas notées  $AB_{r+1}$ . Le polynôme  $P_{n,r}$  interpole  $f$  en les points  $t_n, t_{n-1}, \dots, t_{n-r}$ , où  $r \geq 0$  est fixé, c'est-à-dire l'unique polynôme de degré  $r$  vérifiant

$$P_{n,r}(t_{n-i}) = f(t_{n-i}, y(t_{n-i})), \quad 0 \leq i \leq r.$$

Si l'on écrit ce polynôme dans la base de Lagrange, on a

$$P_{n,r}(t) = \sum_{i=0}^r f(t_{n-i}, y(t_{n-i})) L_{n,i,r}(t),$$

où  $L_{n,i,r}(t) = \prod_{\substack{0 \leq j \leq r \\ j \neq i}} \frac{t - t_{n-j}}{t_{n-i} - t_{n-j}}.$

Par conséquent

$$\begin{aligned} y(t_{n+1}) &= y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt \\ &\simeq y(t_n) + \int_{t_n}^{t_{n+1}} P_{n,r}(t) dt \\ &= y(t_n) + h_n \sum_{i=0}^r b_{n,i,r} f(t_{n-i}, y(t_{n-i})), \end{aligned}$$

avec

$$b_{n,i,r} = \frac{1}{h_n} \int_{t_n}^{t_{n+1}} L_{n,i,r}(t) dt. \quad (4.8)$$

Soit donc le schéma suivant, appelé schéma d'Adams-Bashforth à  $r + 1$  pas

$$y_{n+1} = y_n + h_n \sum_{i=0}^r b_{n,i,r} f(t_{n-i}, y_{n-i}).$$

Voici quelques schémas d'Adams-Bashforth correspondant à quelques valeurs de  $r$ .

-Pour  $r = 0$  : on a

$$\begin{aligned} P_{n,0}(t) &= f(t_n, y(t_n)), \\ \int_{t_n}^{t_{n+1}} P_{n,0}(t) dt &= h_n f(t_n, y(t_n)). \end{aligned}$$

On obtient le schéma

$$\begin{cases} y_{n+1} = y_n + h_n f(t_n, y_n), & n \geq 0 \\ y_0 \text{ donné.} \end{cases}$$

La méthode  $AB_1$  coïncide donc avec la méthode d'Euler explicite.

-Pour  $r = 1$  : l'algorithme d'Adams-Bashforth à 2 pas s'écrit

$$y_{n+1} = y_n + h_n \left( f_n + \frac{h_n}{2h_{n-1}}(f_n - f_{n-1}) \right), \quad n \geq 1$$

où  $f_n := f(t_n, y_n)$ .

-Pour  $r = 2$  : dans le cas où le pas de temps  $h_n = h$  est constant, l'algorithme d'Adams Bashforth à 3 pas s'écrit

$$y_{n+1} = y_n + \frac{h}{12} (23f_n - 16f_{n-1} + 5f_{n-2}), \quad n \geq 2$$

**Remarque 4.11** *L'intérêt des méthodes à pas multiples est qu'elles n'utilisent qu'une seule évaluation de la fonction  $f$  à chaque pas d'intégration. Cela peut s'avérer un gain de temps conséquent par rapport à une méthode de Runge-Kutta d'ordre élevé qui requiert un grand nombre d'évaluations à chaque pas, et ce, surtout lorsque l'évaluation de la fonction est assez coûteuse.*

#### Exemple 4.12

Appliquons le schéma  $(AB_2)$  au problème de Cauchy de l'Exemple 3.4 :

Le schéma  $(AB_2)$  à pas constant s'écrit

$$y_n = y_{n-1} + \frac{h}{2} (3f_{n-1} - f_{n-2}).$$

Le schéma  $(AB_2)$  donne sur cet exemple, avec  $h = 0.1$ ,

$$\begin{aligned} y_n &= y_{n-1} + \frac{h}{2} (3(1 + y_{n-1}) - (1 + y_{n-2})) = (1 + 3\frac{h}{2})y_{n-1} - \frac{h}{2}y_{n-2} + h \\ &= 1.15y_{n-1} - 0.05y_{n-2} + 0.1. \end{aligned}$$

La méthode n'est pas auto-démarrante : le calcul de  $y_1$  doit être effectué par une méthode à un pas. On utilise par exemple, la méthode d'Euler modifiée qui donne  $y_1 = 0.105$ .

Rassemblons dans un tableau les valeurs approchées obtenues, les valeurs exactes et l'erreur effective  $e_n = y(t_n) - y_n$ .

$t_n$	$y_n$	$y(t_n)$	$e_n$
0.0000	0.0000	0.0000	0.0000
0.1	0.105	0.1052	0.0002
0.2	0.2208	0.2214	0.0007
0.3	0.3486	0.3499	0.0012
0.4	0.4899	0.4918	0.0020
0.5	0.6459	0.6487	0.0028
0.6	0.8183	0.8221	0.0038
0.7	1.0088	1.0138	0.0050
0.8	1.2192	1.2255	0.0064
0.9	1.4516	1.4596	0.0080
1.	1.7084	1.7183	0.0099

On voit sur cet exemple que les résultats obtenus par la méthode (RK4) sont meilleurs que ceux obtenus par la méthode (AB<sub>2</sub>).

Dans ce qui suit, on donne deux résultats concernant les coefficients  $b_{n,i,r}$ .

**Proposition 4.13**

i) Lorsque le pas  $h_n = h$  est constant, les coefficients  $b_{n,i,r}$  donnés par la formule (4.8) sont indépendants de  $n$ .

ii) On a toujours  $\sum_{i=0}^r b_{n,i,r} = 1$ .

**Preuve :**

i) En effectuant le changement de variable  $s = \frac{t - t_n}{h}$ , on trouve

$$b_{n,i,r} = (-1)^i \int_0^1 \frac{s(s+1)\dots(s+(i-1))(s+(i+1))\dots(s+r)}{i!(r-i)!} ds, \quad 0 \leq i \leq r,$$

d'où le résultat.

ii) Pour  $f_n = \dots = f_{n-r} = 1$  on a  $p_{n,r}(t) \equiv 1$ , par conséquent

$$\int_0^{t_{n+1}} p_{n,r}(t) dt = h_n = h_n \sum_{i=0}^r b_{n,i,r} \cdot 1.$$

■

On donne le tableau suivant pour les petites valeurs de  $r$ , les coefficients  $b_{n,i,r}$  correspondants sont donnés

$r$	$b_{0,r}$	$b_{1,r}$	$b_{2,r}$	$b_{3,r}$	$\dots$	$\lambda_r := \sum_{i=0}^r  b_{i,r} $
0	1					1
1	$\frac{3}{2}$	$-\frac{1}{2}$				2
2	$\frac{23}{12}$	$-\frac{16}{12}$	$\frac{5}{12}$			$3,66\dots$
3	$\frac{55}{24}$	$-\frac{59}{24}$	$\frac{37}{24}$	$-\frac{9}{24}$		$6,6\dots$

#### 4.2.1 Erreur de consistance et ordre de la méthodes $AB_{r+1}$

Soit  $y$  une solution exacte du problème de Cauchy  $\{(E), (I)\}$ . L'erreur de consistance à l'étape  $n$  est donnée par

$$e_n = y(t_{n+1}) - \bar{y}_{n+1} = y(t_{n+1}) - \left( y(t_n) + \int_{t_n}^{t_{n+1}} P_{n,r}(t) dt \right) = \int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}(t)) dt,$$

où  $P_{n,r}$  est précisément le polynôme d'interpolation de la fonction  $y' = f$  au points  $t_{n-i}$ ,  $0 \leq i \leq r$ . Nous avons le résultat suivant

**Théorème 4.14** La méthode d'Adams Bashforth à  $r+1$  pas est d'ordre  $r+1$ .

**Preuve :** D'après le théorème de la moyenne, il existe un point  $\theta \in ]t_n, t_{n+1}[$  tel que  $e_n = h_n(y'(\theta) - P_{n,r}(\theta))$ . La formule donnant l'erreur d'interpolation (voir annexe) donne :

$$y'(\theta) - P_{n,r}(\theta) = \frac{1}{(r+1)!} y^{(r+2)}(\xi) \pi_{n,r}(\xi),$$

où  $\xi \in ]t_{n-r}, t_{n+1}[$  est un point intermédiaire entre  $\theta$  et les points  $t_{n-i}$ , et

$$\pi_{n,r}(t) = \prod_{0 \leq i \leq r} (t - t_{n-i}).$$

Si  $\xi \in ]t_{n-j}, t_{n-j+1}[$ ,  $0 \leq j \leq r$  on a l'inégalité

$$|\xi - t_{n-i}| \leq (1 + |j - i|)h_{max}, \forall 0 \leq i, j \leq r.$$

D'où

$$\begin{aligned} |\pi_{n,r}(\xi)| &= \prod_{i=0}^j (\xi - t_{n-i}) \prod_{i=j+1}^r (\xi - t_{n-i}) \\ &\leq h_{max}^{r+1} (1+j) \cdots (1+1) 1(1+1) \cdots (1+r-j) \\ &= h_{max}^{r+1} (j+1)!(r-j+1)! \\ &\leq h_{max}^{r+1} (r+1)!, \end{aligned} \tag{4.9}$$

en majorant 2 pas  $j+2, \dots, (r-j+1)$  par  $(r+1)$ . On en déduit par conséquent

$$|y'(\theta) - P_{n,r}(\theta)| \leq |y^{(r+2)}(\xi)| h_{max}^{r+1},$$

ce qui donne la majoration cherchée de l'erreur de consistance :

$$|e_n| \leq |y^{(r+2)}(\xi)| h_n h_{max}^{r+1} \leq C h_n h_{max}^{r+1},$$

avec  $C = \max_{t \in [t_0, t_0+T]} |y^{(r+2)}(t)|$ . L'ordre ne peut être  $\geq r+2$ . Prenons comme contre exemple la fonction  $f(t, y) = t^{r+1}$ , la constante  $C$  est égale dans ce cas à  $(r+1)!$ . La méthode d'Adams Bashforth à  $r+1$  pas est donc d'ordre  $r+1$ . ■

#### 4.2.2 Stabilité de la méthode $AB_{r+1}$

**Théorème 4.15** *On suppose que  $f$  est  $k$ -lipschitzienne par rapport à  $y$ , et que les sommes  $\sum_{i=0}^r |b_{n,i,r}|$  sont majorées indépendamment de  $n$  par une constante  $\lambda_r$ . Alors la méthode  $AB_{r+1}$  est stable avec une constante de stabilité  $S = e^{\lambda_r k T}$ .*

**Preuve :**

Soit  $(\tilde{y}_n)$  la suite récurrente perturbée telle que

$$\begin{cases} \tilde{y}_{n+1} = \tilde{y}_n + h_n \sum_{i=0}^r b_{n,i,r} \tilde{f}_{n-i} + \varepsilon_n, & r \leq n \leq N-1, 0 \leq i \leq r, \\ \tilde{f}_{n-i} = f(t_{n-i}, \tilde{y}_{n-i}) \end{cases}$$

Posons  $\theta_n := \max_{0 \leq i \leq n} |y_i - \tilde{y}_i|$ . On a

$$|f_{n-i} - \tilde{f}_{n-i}| \leq k |y_{n-i} - \tilde{y}_{n-i}| \leq k \theta_n,$$

$$\begin{aligned} |y_{n+1} - \tilde{y}_{n+1}| &\leq |y_n - \tilde{y}_n| + k h_n \sum_{i=0}^r |b_{n,i,r}| \theta_n + |\varepsilon_n|. \\ &\leq (1 + k h_n \lambda_r) \theta_n + |\varepsilon_n|. \end{aligned}$$

Comme  $\theta_{n+1} = \max(\theta_n, |y_{n+1} - \tilde{y}_{n+1}|)$ , on en déduit

$$\theta_{n+1} \leq (1 + kh_n \lambda_r) \theta_n + |\varepsilon_n|.$$

Le lemme de Gronwall implique alors

$$\theta_N \leq \left( \theta_r + \sum_{n=r}^{N-1} |\varepsilon_n| \right) e^{\lambda_r k(t_N - t_r)}.$$

Donc la méthode est stable avec constante de stabilité  $S = e^{\lambda_r kT}$ . ■

**Remarque 4.16** *D'après le dernier tableau on voit que la constante  $\lambda_r$  (quand  $h$  est constant) croît vite avec  $r$ . La stabilité devient donc de moins en moins bonne et c'est l'un des inconvénients les plus sérieux des méthodes d'Adams Bashforth lorsque  $r$  est grand. En pratique, on se limite le plus souvent aux cas  $r = 1$  ou  $2$ .*

### 4.3 Schémas d'Adams-Moulton

Nous allons procéder de la même manière que pour les schémas d'Adams-Bashforth. Cette fois-ci cependant, nous allons approcher  $f$  par l'unique polynôme noté ici  $P_{n,r}^*$  de degré  $r + 1$  vérifiant

$$P_{n,r}^*(t_{n-i}) = f(t_{n-i}, y(t_{n-i})), \quad -1 \leq i \leq r.$$

Dans la base de Lagrange,  $P_{n,r}^*$  s'écrit

$$P_{n,r}^*(t) = \sum_{i=-1}^r f(t_{n-i}, y(t_{n-i})) L_{n,i,r}^*(t)$$

où  $L_{n,i,r}^*$  sont les polynômes de Lagrange associés aux noeuds  $t_{n-i}$ ,  $-1 \leq i \leq n$ , i.e.

$$L_{n,i,r}^*(t) = \prod_{\substack{-1 \leq j \leq r \\ j \neq i}} \frac{t - t_{n-j}}{t_{n-i} - t_{n-j}}.$$

On remarque donc que l'on inclut  $y(t_{n+1})$  (quand  $i = -1$ ) dans les valeurs "connues", c'est pourquoi nous obtenons ainsi une famille de schémas multi-pas **implicites** appelés schémas d'Adams-Moulton à  $r + 1$  pas, notés  $AM_{r+1}$

$$y_{n+1} = y_n + h_n \sum_{i=-1}^r b_{n,i,r}^* f_{n-i}, \quad n \geq r$$

avec

$$b_{n,i,r}^* = \frac{1}{h_n} \int_{t_n}^{t_{n+1}} L_{n,i,r}^*(t) dt.$$

Il faut savoir maintenant comment régler le caractère implicite de cette méthode. Notons  $u_n$  la quantité explicite

$$u_n = y_n + h_n \sum_{i=0}^r b_{n,i,r}^* f_{n-1}.$$

La valeur  $y_{n+1}$  cherchée est donc la solution de l'équation

$$x = u_n + h_n b_{n,-1,r}^* f(t_{n+1}, x).$$

On va donc calculer la suite itérée  $x_{p+1} = \varphi(x_p)$  avec

$$\varphi(x) = u_n + h_n b_{n,-1,r}^* f(t_{n+1}, x).$$

Comme  $\varphi'(x) = h_n b_{n,-1,r}^* f_y'(t_{n+1}, x)$ , et si  $f$  est  $k$ -lipschitzienne en  $y$ , il suffit que  $h_n < \frac{1}{k|b_{n,-1,r}^*|}$  pour que  $\varphi'(x) < 1$  et par conséquent  $\varphi$  est contractante. La solution  $y_{n+1}$  est alors unique d'après le théorème du point fixe et l'algorithme itératif

$$x_{p+1} = u_n + h_n b_{n,-1,r} f(t_{n+1}, x_p).$$

converge vers l'unique point fixe *i.e.*  $y_{n+1}$ . On choisira une valeur initiale  $x_0$  qui soit une approximation de  $y_{n+1}$ , par exemple la valeur donnée par la méthode d'Adams-Baschforth :

$$x_0 = y_n + h_n \sum_{i=0}^r b_{n,i,r}^* f_{n-i}.$$

On arrête l'itération pour  $|x_{p+1} - x_p| \leq \varepsilon$ , où  $\varepsilon$  est la précision donnée, et on prend  $y_{n+1}$  la dernière valeur  $x_p$  calculée.

#### Exemple 4.17

- Pour  $r = 0$  : le polynôme  $P_{n,0}^*$  est le polynôme de degré 1 qui interpole  $f$  en  $t_n$  et  $t_{n+1}$ , soit

$$P_{n,0}^*(t) = f_n + \frac{f_{n+1} - f_n}{h_n}(t - t_n),$$

$$\int_{t_n}^{t_{n+1}} P_{n,0}^*(t) dt = h_n \left( \frac{1}{2} f_{n+1} + \frac{1}{2} f_n \right).$$

On obtient ainsi la méthode dite des trapèzes (ou méthode de Crank-Nicolson) :

$$y_{n+1} = y_n + h_n \left( \frac{1}{2} f_{n+1} + \frac{1}{2} f_n \right)$$

- Pour  $r = 1$  : le polynôme  $P_{n,1}^*$  interpole  $f$  en  $t_{n-1}, t_n$  et  $t_{n+1}$ , soit

$$P_{n,1}^*(t) = f_{n+1} \frac{(t - t_n)(t - t_{n-1})}{h_n(h_n + h_{n-1})} - f_n \frac{(t - t_{n-1})(t - t_{n+1})}{h_n h_{n-1}} + f_{n-1} \frac{(t - t_n)(t - t_{n+1})}{h_{n-1}(h_n + h_{n-1})},$$

et

$$\begin{aligned} y_{n+1} &= y_n + \int_{t_n}^{t_{n+1}} P_{n,1}^*(t) dt \\ &= y_n + h_n \left( \frac{2h_n + 3h_{n-1}}{6(h_n + h_{n-1})} f_{n+1} + \frac{h_n + 3h_{n-1}}{6h_{n-1}} f_n - \frac{h_n^2}{6h_{n-1}(h_n + h_{n-1})} f_{n-1} \right). \end{aligned}$$

Dans le cas où la pas  $h_n = h$  est constant, cette formule se réduit à

$$y_{n+1} = y_n + h \left( \frac{5}{12} f_{n+1} + \frac{8}{12} f_n - \frac{1}{12} f_{n-1} \right).$$



On donne le tableau suivant pour les petites valeurs de  $r$ , les coefficients  $b_{n,i,r}^*$  correspondants sont aussi donnés (l'énoncé de la proposition 4.13 reste valable pour les coefficients  $b_{n,i,r}^*$ ) :

$r$	$b_{-1,r}^*$	$b_{0,r}^*$	$b_{1,r}^*$	$b_{2,r}^*$	$b_{3,r}^*$	$\dots$	$\lambda_r^* = \sum_{i=-1}^r  b_{i,r}^* $
0	$\frac{1}{2}$	$\frac{1}{2}$					1
1	$\frac{5}{12}$	$\frac{8}{12}$	$-\frac{1}{12}$				1, 16...
2	$\frac{9}{24}$	$\frac{19}{24}$	$-\frac{5}{24}$	$\frac{1}{24}$			1, 41...
3	$\frac{251}{720}$	$\frac{646}{720}$	$-\frac{264}{720}$	$\frac{106}{720}$	$-\frac{19}{720}$		1, 78...

#### 4.3.1 Erreur de consistance et ordre de la méthode $AM_{r+1}$

Soit  $y$  une solution exacte du problème de Cauchy  $\{(E), (I)\}$ . Nous avons le résultat suivant :

**Théorème 4.18** *Si  $f$  est  $k$ -lipschitzienne en  $y$ , la méthode d'Adams-Moulton à  $r+1$  pas est d'ordre  $r+2$ .*

**Preuve :** Sous l'hypothèse  $y_{n-i} = y(t_{n-i})$ ,  $0 \leq i \leq n$ , l'erreur de consistance est donnée par

$$\begin{aligned}
e_n &= y(t_{n+1}) - \bar{y}_{n+1} \\
&= y(t_{n+1}) - \left( y(t_n) + h_n \sum_{0 \leq i \leq r} b_{n,i,r}^* f(t_{n-i}, y(t_{n-i})) + h_n b_{n,-1,r}^* f(t_{n+1}, y_{n+1}) \right) \\
&= y(t_{n+1}) - \left( y(t_n) + h_n \sum_{-1 \leq i \leq r} b_{n,i,r}^* f(t_{n-i}, y(t_{n-i})) \right) \\
&\quad + h_n b_{n,-1,r}^* \left( f(t_{n+1}, y(t_{n+1})) - f(t_{n+1}, y_{n+1}) \right). \\
e_n &= \int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}^*(t)) dt + h_n b_{n,-1,r}^* (f(t_{n+1}, y(t_{n+1})) - f(t_{n+1}, y_{n+1})).
\end{aligned}$$

Supposons que  $f(t, y)$  soit lipschitzienne de rapport  $k$  en  $y$ . Alors il vient

$$\begin{aligned}
|e_n| &\leq \left| \int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}^*(t)) dt \right| + h_n b_{n,-1,r}^* k |e_n| \\
&\leq \frac{1}{1 - h_n b_{n,-1,r}^* k} \left| \int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}^*(t)) dt \right|.
\end{aligned}$$

Quand le pas  $h_n$  est suffisamment petit, on écrit

$$|e_n| \leq \left| \int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}^*(t)) dt \right| (1 + O(h_n)).$$

Par ailleurs la formule de la moyenne donne

$$\begin{aligned}
\int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}^*(t)) dt &= h_n (y'(\theta) - P_{n,r}^*(\theta)), \theta \in ]t_n, t_{n+1}[ , \\
y'(\theta) - P_{n,r}^*(\theta) &= \frac{1}{(r+2)!} y^{(r+3)}(\xi) \pi_{n,r}^*(\xi), \xi \in ]t_{n-r}, t_{n+1}[ ,
\end{aligned}$$

où

$$\pi_{n,r}^*(t) = \prod_{-1 \leq i \leq r} (t - t_{n-i}).$$

Il résulte de la majoration (4.9) que

$$\begin{aligned} |\pi_{n,r}^*(\xi)| &= |\xi - t_{n+1}| |\pi_{n,r}(\xi)| \\ &\leq (r+1)h_{\max}(r+1)!h_{\max}^{r+1} \\ &\leq (r+2)!h_{\max}^{r+2}, \end{aligned}$$

d'où

$$\left| \int_{t_n}^{t_{n+1}} (y'(t) - P_{n,r}^*(t)) dt \right| \leq |y^{(r+3)}(\xi)| h_n h_{\max}^{r+2}.$$

Par conséquent l'erreur de consistance admet la majoration

$$|e_n| \leq C h_n h_{\max}^{r+2} (1 + O(h_n)),$$

Avec  $C = \max_{t \in [t_0, t_0+T]} |y^{(r+3)}|$ . Donc la méthode  $AM_{r+1}$  est d'ordre  $r+2$ . ■

**Remarque 4.19** *Il convient d'initialiser les  $r$  premières valeurs  $y_1, \dots, y_r$  à l'aide d'une méthode de Runge Kutta d'ordre  $r+2$ .*

#### 4.3.2 Stabilité de la méthode $AM_{r+1}$

Supposons que  $f(t, y)$  soit  $k$ -lipschitzienne en  $y$ . Notons par

$$\lambda_r^* := \max_{1 \leq n \leq N} \sum_{-1 \leq i \leq r} |b_{n,i,r}^*|$$

et

$$\gamma_r^* := \max_{1 \leq n \leq N} |b_{n,-1,r}^*|$$

D'après le théorème du point fixe, la méthode de résolution itérative pour  $y_{n+1}$  fonctionne dès que  $h_n < \frac{1}{|b_{n,-1,r}^*|k}$ , en particulier dès que  $h_{\max} < \frac{1}{\gamma_r^* k}$ , ce que nous supposons désormais.

Soit  $\tilde{y}_n$  une suite perturbée telle que

$$\begin{cases} \tilde{y}_{n+1} = \tilde{y}_n + h_n(b_{n,-1,r}^* \tilde{f}_{n+1} + \sum_{0 \leq i \leq r} b_{n,i,r}^* \tilde{f}_{n-i}) + \varepsilon_n, \\ \tilde{f}_{n-i} = f(t_{n-i}, \tilde{y}_{n-i}), \quad r \leq n \leq N, \end{cases}$$

Posons  $\theta_n = \max_{0 \leq i \leq n} |\tilde{y}_i - y_i|$ . Comme on a  $\theta_{n+1} = \max(|\tilde{y}_{n+1} - y_{n+1}|, \theta_n)$ , il vient

$$\theta_{n+1} \leq \theta_n + k h_n (|b_{n,-1,r}^*| \theta_{n+1} + \sum_{0 \leq i \leq r} |b_{n,i,r}^*| \theta_n) + |\varepsilon_n|,$$

$$\begin{aligned} \theta_{n+1} (1 - |b_{n,-1,r}^*| k h_n) &\leq \theta_n (1 + \sum_{0 \leq i \leq r} |b_{n,i,r}^*| k h_n) + |\varepsilon_n| \\ &\leq (1 + \sum_{0 \leq i \leq r} |b_{n,i,r}^*| k h_n) (\theta_n + |\varepsilon_n|). \end{aligned}$$

Or  $1 - |b_{n,-1,r}^*|kh_n \geq 1 - \gamma_r^*kh_{max} > 0$ , par suite

$$\begin{aligned}\theta_{n+1} &\leq \frac{1 + \sum_{0 \leq i \leq r} |b_{n,i,r}^*|kh_n}{1 - |b_{n,-1,r}^*|kh_n}(\theta_n + |\varepsilon_n|), \\ \theta_{n+1} &\leq \left(1 + \frac{\sum_{-1 \leq i \leq r} |b_{n,i,r}^*|kh_n}{1 - |b_{n,-1,r}^*|kh_n}\right)(\theta_n + |\varepsilon_n|), \\ \theta_{n+1} &\leq (1 + \Lambda h_n)(\theta_n + |\varepsilon_n|),\end{aligned}$$

avec  $\Lambda = \frac{\lambda_r^*k}{1 - \gamma_r^*kh_{max}}$ .

D'où d'après le lemme de Gronwall :

$$\theta_n \leq e^{\Lambda T}(\theta_r + \sum_{i=r}^n |\varepsilon_i|),$$

et la constante de stabilité  $S = e^{\Lambda T}$ . Lorsque  $h_{max}$  est assez petit devant  $\frac{1}{\gamma_r^*k}$  on obtient :  $S \simeq e^{\lambda_r^*kT}$ .

**Remarque 4.20** *Le tableau précédent montre que la méthode  $AM_{r+1}$  est plus stable que la méthode d' $AB_{r+2}$ .*

## 5 Méthodes de prédiction-correction

Nous avons déjà évoqué les difficultés pratiques rencontrées lors de l'utilisation d'une méthode à pas multiples implicite, liées à la résolution numérique à chaque étape de l'équation (4.2) quand  $\beta_{-1} \neq 0$ , généralement non linéaire, par une méthode des approximations successives. Bien que l'on puisse garantir, que la suite définie par la relation de récurrence (4.4) sera convergente pour toute initialisation arbitraire, on ne sait cependant pas prédire combien d'itération et donc combien d'évaluation de la fonction  $f$  seront nécessaires pour atteindre une précision voulue. Cette incertitude sur le coût de calcul des méthodes implicites rend l'emploi de ces dernières délicat dans certains domaines d'applications. On peut évidemment chercher à rendre cette étape moins coûteuse en fournissant une initialisation raisonnable, mais cela ne permettra pas de contrôler le nombre d'itérations de point fixe réellement effectuées. L'idée des méthodes de prédiction-correction repose sur la prise en compte de cette remarque, en tirant parti d'une méthode explicite : si on prend comme exemple, le schéma (4.2) quand  $\beta_{-1} = 0$ , qualifié de prédicteur, pour obtenir une approximation notée ici  $py_{n+1}$  de la valeur  $y_{n+1}$  solution de l'équation (4.2) recherchée dont on se sert pour effectuer un nombre fixé à l'avance d'itérations de point fixe associées à une méthode implicite, alors appelée le correcteur. La méthode vise à combiner la stabilité de la méthode implicite et la rapidité computationnelle de la méthode explicite. Nous distinguerons le correcteur du prédicteur en attachant des astérisques à tout paramètre s'y rapportant, comme ses coefficients  $\alpha_i^*$  et  $\beta_i^*$ . L'implémentation d'une méthode

de prédiction-correction comporte plusieurs phases, que nous allons maintenant décrire. En supposant que les approximations  $y_{n-r}, f_{n-r}, \dots, y_n, f_n$ , ont été calculées, la prédiction consiste en l'obtention de la valeur  $py_{n+1}$  donnée par

$$py_{n+1} = \sum_{i=0}^r \alpha_i y_{n-i} + h \sum_{i=0}^r \beta_i f_{n-i} \quad (5.1)$$

suit une évaluation de la fonction  $f$  utilisant cette approximation,

$$pf_{n+1} = f(t_{n+1}, py_{n+1}) \quad (5.2)$$

qui permet alors une correction

$$y_{n+1}^{(1)} = h\beta_{-1}^* pf_{n+1} + \sum_{i=0}^r \alpha_i^* y_{n-i} + h \sum_{i=0}^r \beta_i^* f_{n-i} \quad (5.3)$$

Un moyen mnémotechnique pour décrire les diverses mises en œuvre possibles à partir de ces trois phases est de désigner ces dernières par les lettres P, E et C, l'ordre des lettres indiquant leur enchaînement dans la méthode. Par exemple, à chaque étape de la résolution, une méthode de type PEC effectuée, dans cet ordre, les calculs (5.1), (5.2), (5.3), suivis des affectations  $y_{n+1} = y_{n+1}^{(1)}$  et  $f_{n+1} = pf_{n+1}$ . Remarquons qu'on aurait pu choisir d'utiliser la valeur  $y_{n+1}^{(1)}$  pour mettre à jour  $f_{n+1}$  en effectuant une nouvelle évaluation de la fonction  $f$ ,

$$f(t_{n+1}, y_{n+1}^{(1)})$$

ou même d'utiliser cette évaluation pour faire une seconde itération de correction

$$y_{n+1}^{(2)} = h\beta_{-1}^* f(t_{n+1}, y_{n+1}^{(1)}) + \sum_{i=0}^r \alpha_i^* y_{n-i} + h \sum_{i=0}^r \beta_i^* f_{n-i}$$

et alors poser  $y_{n+1} = y_{n+1}^{(2)}$ , donnant ainsi respectivement lieux aux modes PECE et PECEC de la méthode de prédiction-correction. On regroupe l'ensemble des modes ainsi formés par des combinaisons de ces deux procédés sous la notation condensée  $P(EC)^\mu E^{1-\tau}$ , avec  $\mu$  un entier naturel et  $\tau = 0$  ou  $1$ .

## 5.1 Exemples de méthodes de prédiction-correction en mode PECE

- La méthode de Crank-Nicolson est définie par

$$y_{n+1} = y_n + \frac{h_n}{2}(f_n + f_{n+1}).$$

Cet algorithme n'est autre que la méthode  $(AM)_1$ . Si on utilise le schéma d'Euler explicite comme prédicteur pour initialiser cette méthode, on obtient la méthode de  $(RK2)$  avec  $\alpha = 1$  :

$$\begin{cases} P : py_{n+1} = y_n + h_n f_n \\ E : pf_{n+1} = f(t_{n+1}, py_{n+1}) \\ C : y_{n+1} = y_n + h_n(\frac{1}{2}pf_{n+1} + \frac{1}{2}f_n) \\ E : f_{n+1} = f(t_{n+1}, y_{n+1}). \end{cases}$$

• En général, dans les méthodes d'Adams, le prédicteur est une méthode d'Adams-Bashforth d'ordre  $(r + 1)$  et le correcteur utilise une formule d'Adams-Moulton d'ordre  $(r + 2)$ . Par exemple, à l'ordre 4, on emploiera

$$\begin{cases} P : py_{n+1} = y_n + \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}) \\ E : pf_{n+1} = f(t_{n+1}, py_{n+1}) \\ C : y_{n+1} = y_n + \frac{h}{24}(9pf_{n+1} - 19f_n - 5f_{n-1} + f_{n-2}) \\ E : f_{n+1} = f(t_{n+1}, y_{n+1}). \end{cases}$$

## 5.2 Ordre et consistance des méthodes PECE

Soit  $y$  une solution exacte du problème de Cauchy. Soit  $y_{n+1}^*$  la valeur obtenue à l'aide du correcteur  $(AM)_{r+1}$

$$\begin{cases} y_{n+1}^* = y_n + h_n(b_{n,-1,r}^* f_{n+1}^* + \sum_{i=0}^r b_{n,i,r}^* f_{n-i}) \\ f_{n+1}^* = f(t_{n+1}, y_{n+1}^*). \end{cases}$$

L'erreur de consistance prend la forme

$$e_n^* = y(t_{n+1}) - \overline{y}_{n+1}^*$$

où  $\overline{y}_{n+1}^*$  est l'approximation obtenue en insérant la solution exacte dans le schéma, *i.e*

$$y_{n-i} = y(t_{n-i}), \quad 0 \leq i \leq r.$$

Le prédicteur fournit lui aussi une erreur de consistance notée  $pe_n$ , telle que

$$pe_n = y(t_{n+1}) - \overline{py}_{n+1}.$$

L'erreur de consistance de la méthode PECE est notée  $e_n$ , telle que

$$e_n = y(t_{n+1}) - \overline{y}_{n+1}.$$

Ecrivons

$$\begin{aligned} e_n &= (y(t_{n+1}) - \overline{y}_{n+1}^*) + (\overline{y}_{n+1}^* - \overline{y}_{n+1}) \\ &= e_n^* + (\overline{y}_{n+1}^* - \overline{y}_{n+1}). \end{aligned}$$

Mais

$$\overline{y}_{n+1}^* - \overline{y}_{n+1} = h_n b_{n,-1,r}^* (f_{n+1}^* - pf_{n+1}).$$

Supposons que  $f$  est  $k$ -lipschitzienne en  $y$ , on en déduit

$$|\overline{y}_{n+1}^* - \overline{y}_{n+1}| \leq h_n |b_{n,-1,r}^*| k |\overline{y}_{n+1}^* - \overline{py}_{n+1}|.$$

D'autre part

$$\overline{y}_{n+1}^* - \overline{py}_{n+1} = (y(t_{n+1}) - \overline{py}_{n+1}) - (y(t_{n+1}) - \overline{y}_{n+1}^*),$$

ce qui donne

$$|\overline{y}_{n+1}^* - \overline{py}_{n+1}| \leq |pe_n| + |e_n^*|.$$

Finalement

$$\begin{aligned} |\overline{y}_{n+1}^* - \overline{y}_{n+1}| &\leq kh_n |b_{n,-1,r}^*| (|pe_n| + |e_n^*|), \\ |e_n| &\leq |e_n^*| + |\overline{y}_{n+1}^* - \overline{y}_{n+1}| \leq (1 + kh_n |b_{n,-1,r}^*|) |e_n^*| + kh_n |b_{n,-1,r}^*| |pe_n|. \end{aligned}$$

On voit donc que l'influence du prédicteur est nettement moindre que celle du correcteur puisque son erreur de consistance est en facteur d'un terme  $O(h_n)$ . Le correcteur  $(AM)_{r+1}$  étant d'ordre  $r+2$ , il convient donc de choisir un prédicteur d'ordre  $r+1$ . L'ordre de PECE est donc  $r+2$  dans ce cas.

### 5.3 Stabilité des méthodes PECE

Supposons que le prédicteur soit de la forme

$$py_{n+1} = \sum_{i=0}^r \alpha_{n,i} y_{n-i} + h_n \sum_{i=0}^r \beta_{n,i} f_{n-i}$$

et notons

$$A = \max_n \sum_i |\alpha_{n,i}|, \quad B = \max_n \sum_i |\beta_{n,i}|.$$

On prend la méthode  $(AM)_{r+1}$  comme correcteur.

Soit  $(\tilde{y}_n)$  une suite perturbée telle que

$$\begin{cases} p\tilde{y}_{n+1} = \sum_{i=0}^r \alpha_{n,i} \tilde{y}_{n-i} + h_n \sum_{i=0}^r \beta_{n,i} \tilde{f}_{n-i} \\ \tilde{y}_{n+1} = \tilde{y}_n + h_n (b_{n,-1,r}^* p\tilde{f}_{n+1} + \sum_{i=0}^r b_{n,i,r}^* \tilde{f}_{n-i}) + \varepsilon_n \end{cases}$$

Dans la réalité il s'introduit également une perturbation  $p\varepsilon_n$  au niveau de la prédiction, mais pour notre calcul il sera plus simple de comptabiliser cette erreur dans  $\varepsilon_n$ .

Posons  $\theta_n = \max_{0 \leq i \leq n} |\tilde{y}_i - y_i|$  et  $p\theta_n = |p\tilde{y}_n - py_n|$ . Comme  $f(t, y)$  est supposée  $k$ -lipschitzienne en  $y$ , il vient

$$\begin{cases} p\theta_{n+1} \leq A\theta_n + h_n Bk\theta_n \\ \theta_{n+1} \leq \theta_n + kh_n (|b_{n,-1,r}^*| p\theta_{n+1} + \sum_{i=0}^r |b_{n,i,r}^*| \theta_n) + |\varepsilon_n|. \end{cases}$$

En substituant  $p\theta_{n+1} \leq \theta_n + \theta_n(A - 1 + h_n Bk)$  dans la deuxième ligne, il vient

$$\theta_{n+1} \leq \theta_n \left( 1 + \left( \sum_{i=-1}^r |b_{n,i,r}^*| + |b_{n,-1,r}^*| (A - 1 + h_n Bk) \right) kh_n \right) + |\varepsilon_n|,$$

$$\theta_{n+1} \leq \theta_n (1 + \Lambda h_n) + |\varepsilon_n|,$$

avec  $\Lambda = (\lambda_r^* + \gamma_r^* (A - 1 + Bkh_{\max}))k$ . Le lemme de Gronwall montre donc que la méthode PECE est stable, avec constante de stabilité

$$S = e^{\Lambda T}.$$

Si  $Bkh_{\max} \ll 1$ , on écrit

$$S \simeq e^{(\lambda_r^* + \gamma_r^* (A-1))kT},$$

*i.e.*, la stabilité du prédicteur n'influe pas la stabilité de la méthode PECE, seule la valeur de la constante  $A$  peut influencer sur cette stabilité. Par exemple dans le cas des méthodes  $(AB)_{r+1}$  et quand les coefficients  $b_{n,i,r}$  sont positifs,  $A = 1$  (voir proposition 4.13), par conséquent la constante de stabilité

$$S = e^{\lambda_r^* k T}$$

sera peu différente de celle de la méthode d'Adams-Moulton. On obtient donc des méthodes assez stables, d'un coût modéré en temps de calcul.

## 6 Mise en œuvre

On va montrer dans cette section comment mettre en œuvre quelques schémas de résolution numérique des équations différentielles vus précédemment, à l'aide de MATLAB. Nous renvoyons le lecteur intéressé à [7] pour une description complète du langage de MATLAB.

On considère le problème de Cauchy suivant

$$\begin{cases} y' = t - ty \\ y(0) = 2 \end{cases} \quad (6.1)$$

On utilise cet exemple comme référence pour analyser la précision des différentes solutions numériques. La solution exacte est donnée par  $y(t) = 1 + e^{-\frac{t^2}{2}}$ .

La fonction MATLAB "f.m" correspondant à la donnée  $f$  est la suivante

```
function yp=f(t,y)
yp=t-t*y;
```

Et la fonction MATLAB "fex.m" correspondant à la solution exacte est :

```
function y=fex(t)
y=1+exp(-t.^2/2);
```

### 6.1 La méthode d'Euler explicite

Soit la fonction MATLAB "eulerexplicite.m"; elle renvoie les temps successifs où sont effectuées les solutions approchées.

```
function [liste_t,liste_y]=eulerexplicite(y0,N,T) y=y0;liste_y=[y0];
t=0;liste_t=[0]; h=T/N;
for i=1:N
    y=y+h*f(t,y);
    t=t+h;
    liste_y=[liste_y,y];
    liste_t=[liste_t,t];
end
```

Si l'on souhaite avoir accès aux deux arguments de sortie c.à.d. les temps et les valeurs des approximations, il faut effectuer l'appel comme suit :

```
>> [tps,sol]=eulerexplicite(2,5,2)
tps =
      0      0.4000      0.8000      1.2000      1.6000      2.0000
sol =
  2.0000  2.0000  1.8400  1.5712  1.2970  1.1069
```

## 6.2 La méthode d'Euler implicite

La fonction correspondante est la suivante  
`function[t,y]=eulerimplicite(f,T,y0,h)`

```
N=T/h; t(1)=0; y(1)=y0; fori=1:N
    t(i+1)=t(i)+h;
    ynew=y(i)+h*feval(f,t(i),y(i));
    y(i+1)=y(i)+h*feval(f,t(i+1),ynew);
end
```

Le programme principal suivant fournit les solutions approchées du problème (6.1) à partir de la méthode d'Euler explicite et d'Euler implicite ainsi que la solution exacte. La commande "plot" trace les trajectoires en fonctions du temps.

```
y0=2; t=[0:0.4:2]
[yexpl]=eulerexplicite(2,5,2)
[yimpl]=eulerimplicite('f',2,2,0.4)
yexact=[fex(t)]
%graphique de comparaison entre la solution exacte et la
%la solution approchée par les méthodes
%d'Euler explicite et d'Euler implicite
plot(t,yexpl,'r:',t,yimpl,'b',t,fex(t),'b--','Linewidth',2)
xlabel('temps t','fontsize',14) ylabel('y(t)','fontsize',14)
legend('Euler explicite','Euler implicite','solution exacte')
```

L'exécution de ce programme donne le résultat suivant :

```
t = 0      0.4000      0.8000      1.2000      1.6000      2.0000

yexpl = 2.0000      2.0000      1.8400      1.5712      1.2970      1.1069

yimpl = 2.0000      1.8400      1.6142      1.4137      1.2760      1.1965

yexact = 2.0000      1.9231      1.7261      1.4868      1.2780      1.1353
```



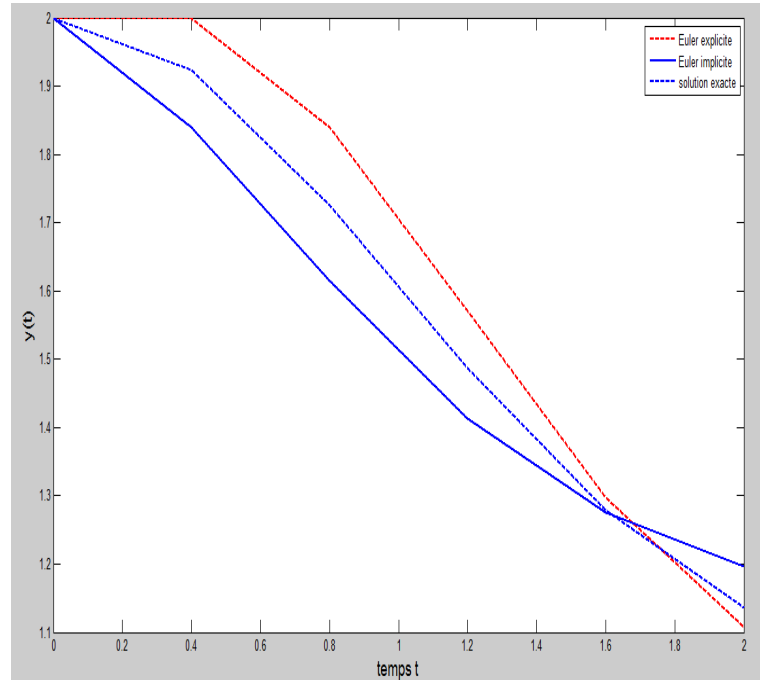


FIG. 1 – Graphique de comparaison entre la solution exacte et la solution approchée par les méthodes d'Euler explicite et d'Euler implicite

### 6.3 La méthode de Runge Kutta 4

```
function[t,y]=rk4(f,T,y0,h) N=T/h;t(1)=0; y(1)=y0;
for i=1:N
    f1=h*feval(f,t(i),y(i));
    f2=h*feval(f,t(i)+h/2,y(i)+f1/2);
    f3=h*feval(f,t(i)+h/2,y(i)+f2/2);
    f4=h*feval(f,t(i)+h,y(i)+f3);
    y(i+1)=y(i)+(f1+2*f2+2*f3+f4)/6;
    t(i+1)=t(i)+h;
end
```

end

Voici le programme de comparaison entre les solutions approchées du problème (6.1) à partir de la méthode d'Euler explicite et de Runge Kutta d'ordre 4 et la solution exacte :

```
y0=2;
N=5;
T=2;
t=[0:0.4:2]
yexpl=eulerexplicite(2,5,2)
yrk4=rk4('f',2,2,0.4)
yexact=[fex(t)]
%graphique de comparaison entre la solution exacte la solution
%approchée par les méthodes d'Euler explicite et de Runge Kutta 4
plot(t,yexpl,'g',t,yrk4,'b',t,fex(t),'r--','Linewidth',2)
```

```
xlabel('temps t','fontsize',14) ylabel('y(t)','fontsize',14)
legend('Euler explicite','Runge Kutta 4','solution exacte')
```

L'exécution de ce programme donne le résultat suivant :

```
t = 0      0.4000      0.8000      1.2000      1.6000      2.0000
```

```
yexpl = 2.0000      2.0000      1.8400      1.5712      1.2970      1.1069
```

```
yrk4 = 2.0000      1.9231      1.7261      1.4868      1.2782      1.1358
```

```
yexact = 2.0000      1.9231      1.7261      1.4868      1.2780      1.1353
```

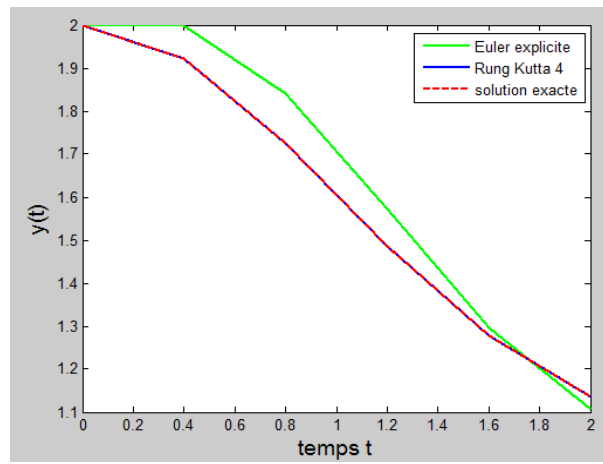


FIG. 2 – Graphique de comparaison entre la solution exacte et la solution approchée par les méthodes d'Euler explicite et de Runge Kutta 4

## 6.4 La méthode d'Adams Bashforth d'ordre 3

Le programme de cette méthode est le suivant :

```
t(1)=0; y(1)=2; h=0.4;
for i=1:6
    f(i)=(t(i)-t(i)*y(i));
    if(i<3)
        y(i+1)=y(i)+h*(t(i)-t(i)*y(i));
    else
        y(i+1)=y(i)+h/12*(23*f(i)-16*f(i-1)+5*f(i-2));
    end
    t(i+1)=t(i)+h;
end
t=[t(1),t(2),t(3),t(4),t(5),t(6)]
y=[y(1),y(2),y(3),y(4),y(5),y(6)]
yexact=[fex(t)]
%graphique de comparaison entre la solution exacte
%et la solution approchée par la méthode d'AB3
```

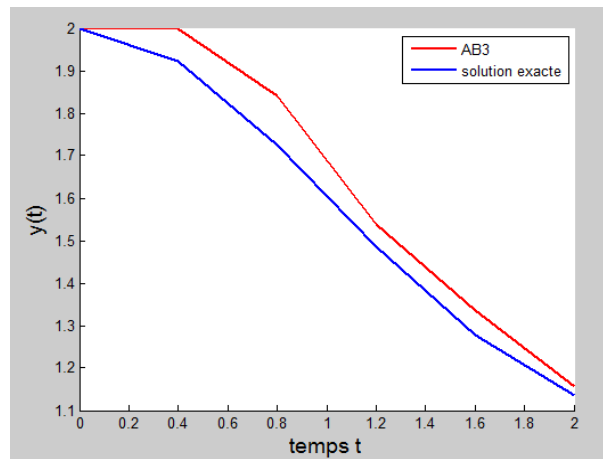
```
hold on plot(t,y,'r',t,fex(t),'b','Linewidth',2)
legend('AB3','solution exacte') xlabel('temps t','fontsize',14)
ylabel('y(t)','fontsize',14)
```

Ce programme donne le résultat suivant :

```
t = 0      0.4000      0.8000      1.2000      1.6000      2.0000
```

```
y = 2.0000  2.0000      1.8400      1.5381      1.3348      1.1565
```

```
yexact = 2.0000      1.9231      1.7261      1.4868      1.2780      1.1353
```



## 6.5 La méthode de prédiction correction

On prend le prédicteur Euler explicite et le correcteur  $AM_1$  (d'ordre 2), le programme de cette méthode est le suivant

```
function[t,y]=pc(f,T,y0,h) N=T/h; y(1)=y0; t(1)=0;
for i=1:N
    k1=feval(f,t(i),y(i));
    t(i+1)=t(i)+h;
    k2=feval(f,t(i+1),y(i)+(h*k1));
    y(i+1)=y(i)+h*(k1+k2)/2;
end
```

Le programme de comparaison avec la solution exacte est donné par :

```
y0=2; N=5; T=2; [t,ypc]=pc('f',2,2,0.4)
yexact=[fex(t)]
%graphique de comparaison entre la solution exacte
%et la solution approchée par la méthodes de
%prédiction correction
plot(t,ypc,'r--',t,fex(t),'b','Linewidth',2)
xlabel('t','fontsize',14) ylabel('y(t)','fontsize',14)
legend('prédiction correction','solution exacte')
```

L'exécution de ce programme donne le résultat suivant :

t =	0	0.4000	0.8000	1.2000	1.6000	2.0000
ypc =	2.0000	1.9200	1.7228	1.4892	1.2904	1.1556
yexact =	2.0000	1.9231	1.7261	1.4868	1.2780	1.1353

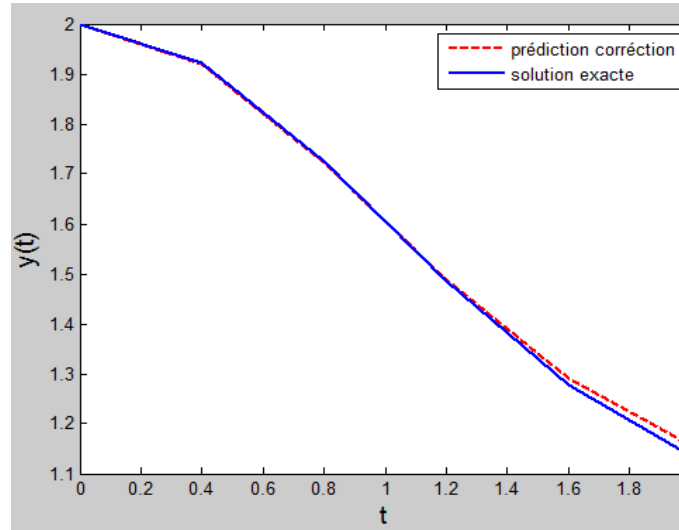


FIG. 4 – Graphique de comparaison entre la solution exacte et la solution approchée par la méthode de prédiction correction

## 6.6 La résolution numérique d'un système d'EDO : Modèle proie/prédateur

**Présentation du modèle :** Le modèle se comporte de la façon suivante :

Lorsque la population des proies augmente, celle des prédateurs peut augmenter car elle a plus de proies pour se nourrir ; inversement lorsque la population des prédateurs devient trop importante, elle se nourrit plus vite que ne peut renouveler la population de proies, conduisant au décroissement des deux populations.

Les populations de deux espèces sont notées  $x$  et  $y$ , les paramètres de leur lois d'évolution naturelle pour les proies et les prédateurs sont respectivement les coefficients  $\mu_1$  et  $\mu_3$ , et les paramètres  $\mu_2$  et  $\mu_4$ , sont les coefficients définissant les influences inter-populations. Le système qui régit ce comportement est le suivant :

$$\begin{cases} x'(t) = x(t)(\mu_1 - \mu_2 y(t)) \\ y'(t) = y(t)(-\mu_3 + \mu_4 x(t)) \end{cases}$$

**La résolution numérique :** Appliquons la méthode d'Euler explicite déjà programmée avec  $\mu_1 = 1.5$ ,  $\mu_2 = 1$ ,  $\mu_3 = 3$ ,  $\mu_4 = 1$ , et des populations initiales de 8 et 4. La méthode d'Euler est indépendante de la dimension. Pour résoudre ce système il suffit donc de redéfinir la fonction f.m comme suit :

```

function yp=f(t,y) yp=[y(1)*(3-y(2));y(2)*(-2+2*y(1))]; end
Puis on appelle la fonction eulerexplicite déjà définie dans la section 6.1 comme suit :
T=10;
N=5000;
y0=[8;4] [tps,sol]=eulerexplicite(y0,N,T);
plot(tps,sol(1,:),tps,sol(2,:)) title('Evolution des populations de
proie prédateur pour le problème Lotka Volterra avec les paramètres
mu=(1.5,1,3,1)) legend('proie','prédateur')

```

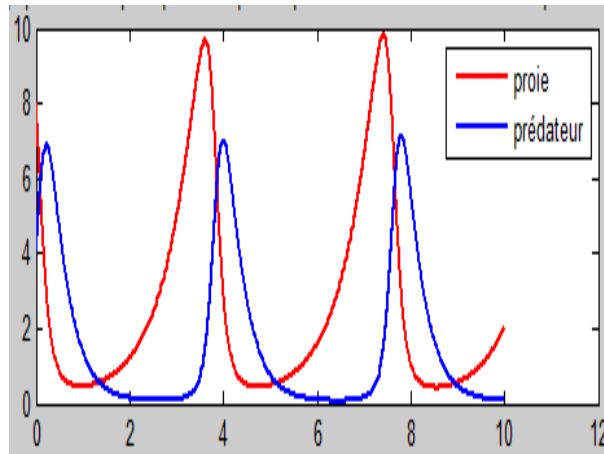


FIG. 5 – Evolution des populations de proie prédateur pour le problème Lotka Volterra avec les paramètres  $\mu = (1.5, 1, 3, 1)$

## 7 Exercices

**7.1** Utiliser la méthode d'Euler pour calculer les premières 4 valeurs de la solution  $y(t)$  de l'équation différentielle

$$y' = \frac{y-t}{y+t}$$

qui satisfait à la condition initiale  $y(0) = 1$ , en prenant le pas  $h = 0.1$ .

**7.2** Trouver la solution approchée au point  $t = 2.1$  du problème de Cauchy

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

en utilisant la méthode de Taylor d'ordre 4 avec  $h = 0.1$ .

**7.3** Soit à résoudre

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

1. Trouver la valeur de  $y$  au points  $t_1 = 0.5$  et  $t_2 = 1.0$  par la méthode d'Euler avec un pas  $h = 0.5$ .
2. Calculer  $y$  au point  $t = 1$  en utilisant successivement les méthodes de Runge-Kutta d'ordre 2 et 4.
3. Quelle est la solution exacte du problème ?.
4. Comparer les valeurs approchées calculées en 1. et 2. avec la valeur exacte correspondante. Conclure.

**7.4** Soit le problème de Cauchy

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

1. En utilisant la méthode de Runge-Kutta d'ordre 4, calculer les premières 2 valeurs approchées de la solution du problème en prenant le pas  $h = 0.2$ .
2. Utiliser les valeurs ainsi obtenues pour calculer la valeur de la solution au point  $t = 2$  par la méthode d'Adams-Baschforth d'ordre 3.
3. Prendre comme prédicteur les valeurs obtenues à la question précédente pour calculer la valeur de la solution au point  $t = 2$  par la méthode d'Adams-Moulton d'ordre 3.
4. Estimer les résultats obtenus si la solution exacte est  $y = \frac{t}{2} - \frac{1}{2t}$ .

**7.5** Soit à résoudre numériquement un problème de Cauchy

$$y' = f(t, y), y(t_0) = y_0$$

où  $f$  est de classe  $C^2$  sur  $[t_0, t_0 + T] \times \mathbb{R}$ . On se donne une subdivision

$$t_0 < t_1 < \dots < t_N = t_0 + T$$

de l'intervalle, et on étudie la méthode numérique suivante : si  $y_n$  est la valeur approchée de la solution au temps  $t_n$  et si  $f_n = f(t_n, y_n)$  on pose

$$(M) \quad y_{n+1} = y_n + \int_{t_n}^{t_{n+1}} p_n(t) dt$$

où  $p_n$  est le polynôme d'interpolation des pentes  $f_n, f_{n-1}$  aux temps  $t_n, t_{n-1}$ . On note  $h_n = t_{n+1} - t_n$ .

1. Calculer explicitement  $y_{n+1}$ . Quelle méthode obtient-on lorsque le pas  $h_n = h$  est constant ?.
2. Soit  $z$  une solution exacte de l'équation différentielle. Déterminer un équivalent de l'erreur de consistance  $e_n$  attachée à la méthode  $(M)$ . Quel est l'ordre de cette méthode ?.
3. Soit une suite  $\tilde{y}_n$  vérifiant la relation de récurrence

$$\tilde{y}_{n+1} = \tilde{y}_n + \int_{t_n}^{t_{n+1}} \tilde{p}_n(t) dt + \varepsilon_n$$

où  $\tilde{p}_n$  interpole les valeurs  $\tilde{f}_n = f(t_n, \tilde{y}_n)$  et  $\tilde{y}_{n-1}$ . La quantité  $\varepsilon_n$  l'erreur commise à l'étape  $n$ . On suppose que  $f(t, y)$  est  $k$  lipschitzienne en  $y$  et on note  $\theta_n = \max_{0 \leq i \leq n} |\tilde{y}_i - y_i|$ .

i) Montrer que

$$\theta_{n+1} \leq \left(1 + kh_n \left(1 + \max\left\{\frac{h_n}{h_{n-1}}, \frac{h_{n-1}}{h_n}\right\}\right)\right) \theta_n + \varepsilon_n.$$

ii) On suppose que le rapport de deux pas consécutifs est majoré par une constante  $\delta$  (avec disons  $1 \leq \delta \leq 2$ ). Etudier la stabilité de la méthode  $(M)$ .

**7.6** On considère la méthode générale à  $r$  pas, définie par : pour  $n \geq r - 1$

$$(\rho, \sigma) \quad y_{n+1} = \sum_{i=1}^r a_i y_{n+1-i} + h \sum_{i=0}^r b_i f(t_{n+1-i}, y_{n+1-i}),$$

avec  $t_n = t_0 + nh$ ; les valeurs de démarrage  $y_0, y_1, \dots, y_{r-1}$  sont données. A cette méthode on associe les polynômes  $\rho$  et  $\sigma$  à coefficients réels définis par

$$\rho(\xi) = \xi^r - \sum_{i=1}^r a_i \xi^{r-i}, \quad \sigma(\xi) = \sum_{i=0}^r b_i \xi^{r-i}.$$

Remarquons que la méthode est entièrement caractérisée par la donnée des polynômes de degré  $r$ ,  $\rho$  et  $\sigma$ .

1. Donner des conditions suffisantes assurant l'existence et l'unicité d'une solution au schéma  $(\rho, \sigma)$ .
2. Montrer que la méthode  $(\rho, \sigma)$  est consistante si, et seulement si,

$$\rho(1) = 0 \text{ et } \rho'(1) = \sigma(1).$$

3. On considère le développement en série entière :

$$\rho(e^z) - z\sigma(e^z) = c_0 + c_1z + \dots + c_mz^m + \dots$$

Montrer que les deux conditions suivantes sont équivalentes :

*i)* : La méthode  $(\rho, \sigma)$  est d'ordre  $k$ ; *ii)* :  $c_0 = c_1 = \dots = c_k = 0$ .

4. Définir, par analogie avec le cours, la stabilité du schéma  $(\rho, \sigma)$  et montrer que, si la fonction  $f$  est uniformément lipschitzienne en  $y$ , ce schéma est stable si, et seulement si, la condition dite de Dahlquist  $(D_0)$  est satisfaite :

$$(D_0) \quad \begin{cases} \text{les racines de } \rho \text{ sont de module } \leq 1 \text{ et} \\ \text{les racines de module } 1 \text{ sont simples.} \end{cases}$$



## Annexe

### Théorème 1 (théorème du point fixe)

Soit  $(E, d)$  un espace métrique complet non vide, et  $f : E \rightarrow E$  une contraction stricte, i.e.  $f$  lipschitzienne de rapport  $L < 1$ . Alors  $f$  admet un unique point fixe  $x \in E$ . De plus, la suite  $(x_n)_{n \in \mathbb{N}}$  définie par  $x_{n+1} = f(x_n)$  converge lorsque  $n$  tend vers l'infini pour toute initialisation  $x_0$  vers ce point fixe.

**Définition 2** Une méthode d'intégration numérique est dite d'ordre  $N$  si la formule approchée est exacte pour tout polynôme de degré  $\leq N$  et inexacte pour au moins un polynôme de degré  $N + 1$ .

### Théorème 3 (formule d'erreur d'interpolation)

Soit  $f : [a, b] \rightarrow \mathbb{R}$  une fonction  $n + 1$  fois dérivable sur  $[a, b]$ . On se donne  $n + 1$  points  $x_0, x_1, \dots, x_n$  dans  $[a, b]$ , deux à deux distincts. Soit  $p_n$  le polynôme d'interpolation de Lagrange de  $f$  aux points  $x_i, i = 0, \dots, n$ . Alors pour tout  $x \in [a, b]$ , il existe un point  $\xi_x \in ]\min(x, x_i), \max(x, x_i)[$  tel que

$$f(x) - p_n(x) = \frac{1}{(n+1)!} \prod_{i=0}^n (x - x_i) f^{(n+1)}(\xi_x).$$

### Théorème 4 (théorème de la moyenne)

Si  $f$  est intégrable sur  $[a, b]$  et si l'on pose  $m = \inf_{x \in [a, b]} f(x)$  et  $M = \sup_{x \in [a, b]} f(x)$  alors

$$m(b-a) \leq \int_a^b f(t) dt \leq M(b-a).$$

Si  $f$  est continue sur  $[a, b]$ , alors il existe  $c \in ]a, b[$  tel que

$$\frac{1}{b-a} \int_a^b f(t) dt = f(c).$$

## Références

- [1] F. Ceschino, J. Kuntzmann, *Problèmes différentiels de conditions initiales*, Dunod, 1963.
- [2] M. Crouzeix, A. Mignot, *Exercices d'analyse numérique des équations différentielles*, Masson, 1993.
- [3] J.-P. Demailly, *Analyse numérique et équations différentielles*, Presse universitaire de Grenoble, 1996.
- [4] A. Fortin, *Analyse numérique pour ingénieurs*, presses internationales, polytechnique, 2001.
- [5] F. Jedrzejewski, *Introduction aux méthodes numériques*, Springer-Verlag, Paris, 2005.
- [6] M. Lakrib, *Cours d'analyse numérique*, OPU, Alger, 2005.
- [7] A. Mesbah, M. Mokhtari, *Apprendre et maîtriser MATLAB*, Springer, 1997.
- [8] A. Quarteroni, R. Sacco, F. Saleri, *Méthodes numériques pour le calcul scientifique, programmes en matlab*, Springer-Verlag, 2000.
- [9] M. SIBONY, J.C. MARDON, *Analyse numérique 2, Approximations et équations différentielles*, Hermann, éditeurs des sciences et des arts, 1982.
- [10] Sites internet :  
  
[https ://www.ceremade.dauphine.fr/ legendre/](https://www.ceremade.dauphine.fr/legendre/),  
  
[math.univ-angers.fr/ granger/anatum/ChapitreV.pdf](http://math.univ-angers.fr/granger/anatum/ChapitreV.pdf),  
  
[http ://moodle.utc.fr/file.php/665/MT09-ch7.pdf](http://moodle.utc.fr/file.php/665/MT09-ch7.pdf),  
  
[lmah.univ-lehavre.fr/ alaoui//livre-MN-Java.pdf](http://lmah.univ-lehavre.fr/alaoui/livre-MN-Java.pdf),  
  
[http ://faccanoni.univ-tln.fr/enseignements.html](http://faccanoni.univ-tln.fr/enseignements.html).