Licence à distance

Chapitre V : Equations différentielles. Méthodes numériques à un pas.

M. Granger

Table des matières

1 Rappels sur le cours d'équations différentielles

Les deux premières sous-sections sont des rappels que nous donnons pour fixer les notations de notions et de résultats développés dans le cours d'équations différentielles : cylindre de sécurité et théorème de Cauchy-Lipschitz. Les deux suivantes sont plus spécifiques à notre propos : régularité des solutions, avec l'introduction des fonctions $f^{[k]}(x,y)$ outil de calcul des dérivées successives d'une solution et méthode d'Euler, qui est la méthode "de base" pour le calcul approché des solutions.

1.1 Généralités

Les données du problème sont les suivantes :
– $U \subset \mathbb{R} \times \mathbb{R}^m$, un ouvert
– $f : U \to \mathbb{R}^m$, une application (au moins continue.)
On cherche les solutions de l'équation différentielle

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

c'est à dire les couples (I,φ), où $I \subset \mathbb{R}$ est un intervalle et $\varphi : I \to \mathbb{R}^m$ est une application dérivable, telle que :

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

On s'intéresse au problème de Cauchy de condition initiale (t_0,y_0), c'est à dire à l'existence d'une solution de (E), sur un intervalle I telle que $t_0 \in I$, et $\varphi(t_0) = y_0$, et à l'éventuelle unicité, pour I fixé, d'une telle solution.

La condition sur φ équivaut à l'équation intégrale :

$$\varphi(t) = y_0 + \int_{t_0}^{t} f(t,\varphi(u))du$$

Dans tout ce qui suit f est au minimum supposée continue.

Cylindres 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), et 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 $\mathbb{R}^m$ est arbitraire.

Proposition 1 .- *Soit $T = min(T_0, \frac{r_0}{M})$. Toute solution au problème de Cauchy pour les conditions initiales (t_0,y_0) satisfait à :*

$$|t - t_0| \leq T \Longrightarrow \|\varphi(t) - y_0\| \leq r_0$$

On appelle $[t_0 - T, t_0 + T] \times \overline{B}(y_0, r_0)$ un *cylindre de sécurité* pour la condition initiale (t_0,y_0).

1.2 Enoncé du théorème de Cauchy-Lipschitz

(sans démonstrations)

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.

Théorème 1 .- *On suppose que $f : U \to \mathbb{R}^m$ est continue et localement lipschitzienne dans la deuxième variable, ce qui signifie que tout point de U admet un voisinage de la forme $J \times V$, tel qu'il existe une constante k (dépendant de $J \times V$!), avec la condition :*

$$\forall t \in J,\ \forall y_1, y_2 \in V,\ \ \|f(t,y_1) - f(t,y_2)\| \leq k\|y_1 - y_2\|$$

Alors, pour tout cylindre de sécurité $C = [t_0 - T, t_0 + T] \times B(y_0, r_0)$, l'équation (E) admet une unique *solution, de condition initiale (t_0,y_0), définie sur l'intervalle $I = [t_0 - T, t_0 + T]$.*

1.3 Sur la régularité des solutions.

Si la fonction f est différentiable de classe $\mathcal{C}^k$, avec $k \geq 1$, on peut en déduire un résultat sur l'ordre de dérivabilité des solutions φ, et calculer les dérivées successives de φ à l'aide de fonctions construites récursivement à partir de f. C'est utile pour obtenir des majorations, fondées sur la formule de Taylor, de l'erreur commise dans une approximation numérique de φ.

Proposition 2 *.- 1) Si la fonction f est de classe $\mathcal{C}^k$, avec $k \geq 1$, toute solution $t \mapsto z(t)$, de l'équation (E) est de classe $\mathcal{C}^{k+1}$.*
2) Les fonctions $f^{[i]} : U \to \mathbb{R}$, définies récursivement, pour $i = 0, \cdots, k$, par les formules :

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

sont de classes respectives $\mathcal{C}^{k-i}$, et pour toute solution $z(t)$ de l'équation (E), ses dérivées successives s'obtiennent par les formules :

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

Démonstration.- On démontre ce résultat par récurrence sur i.

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

-Pour la différentiabilité de z, on montre par récurrence sur i, pour $0 \leq i \leq k$, l'existence de $z^{(i+1)}$, et la formule annoncée en 2).
Pour $i = 0$, il s'agit simplement de l'égalité $z'(t) = f(t, z(t))$, qui signifie que $t \to z(t)$ est une solution. On voit sur cette relation que z' est continûment dérivable et la formule de dérivation des fonctions composées s'écrit :

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

Pour le pas général de la récurrence, on suppose que $z^{(i+1)}(t) = f^{[i]}(t, z(t))$, avec $1 \leq i < k$. Alors puisque $f^{[i]}$ est de classe $\mathcal{C}^{k-i}$, avec $k - i > 0$, on en déduit que $z^{(i+1)}$ est encore continûment dérivable et par un calcul semblable au cas de z'', on trouve :

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

□

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

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

Exercice : On suppose que f est de classe $\mathcal{C}^1$, déterminer par une inégalité l'ensemble des points (t_0, y_0) au voisinage desquels une solution $z(t)$ telle que $z(t_0) = y_0$ soit convexe. Expliciter le résultat pour l'équation $y' = t - y^2$.

1.4 Méthode d'Euler.

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

Dans toute la suite, on gardera les notations suivantes :

$$t_0 < t_1 < \cdots < t_N = t_0 + T, \quad h_n = t_{n+1} - t_n \ \text{ pour } \ n = 0, \cdots, N-1$$

définit une subdivision de $[t_0, t_0 + T]$, et h_n est appelé le $(n+1)$ième pas de la subdivision. On parle de subdivision à pas constant si $t_n = t_0 + \frac{nT}{N}$, $h_n = \frac{T}{N}$ indépendant de n.

La méthode d'Euler ou méthode de la tangente consiste à approcher une solution (éventuelle !) $z(t)$ de (E) par une fonction $y(t)$ continue affine par morceaux, affine en restriction à chaque intervalle $[t_n, t_{n+1}]$, de pente $f(t_n, y_n)$ où $y_n = y(t_n)$. Ceci donne le calcul par récurrence suivant des y_n appelé *schéma d'Euler* et l'expression de $y(t)$:

$$\begin{cases} y(t_0) = y_0 & \text{condition initiale ,} \\ y_{n+1} = y_n + h_n f(t_n, y_n) & \text{formule de récurrence pour les } y_n, \\ y(t) = y_n + (t - t_n) f(t_n, y_n) & \text{si } t \in [t_n, t_{n+1}] \ . \end{cases}$$

Ce schéma peut être illustré par le dessin suivant où on a représenté simultanément une solution exacte $z(t)$ passant par les point M_i de coordonnées $(t_i, z(t_i))$, et une solution approchée linéaire par morceaux de graphe la ligne brisée $(A_0, A_1 \cdots)$, A_i de coordonnées (t_i, y_i).

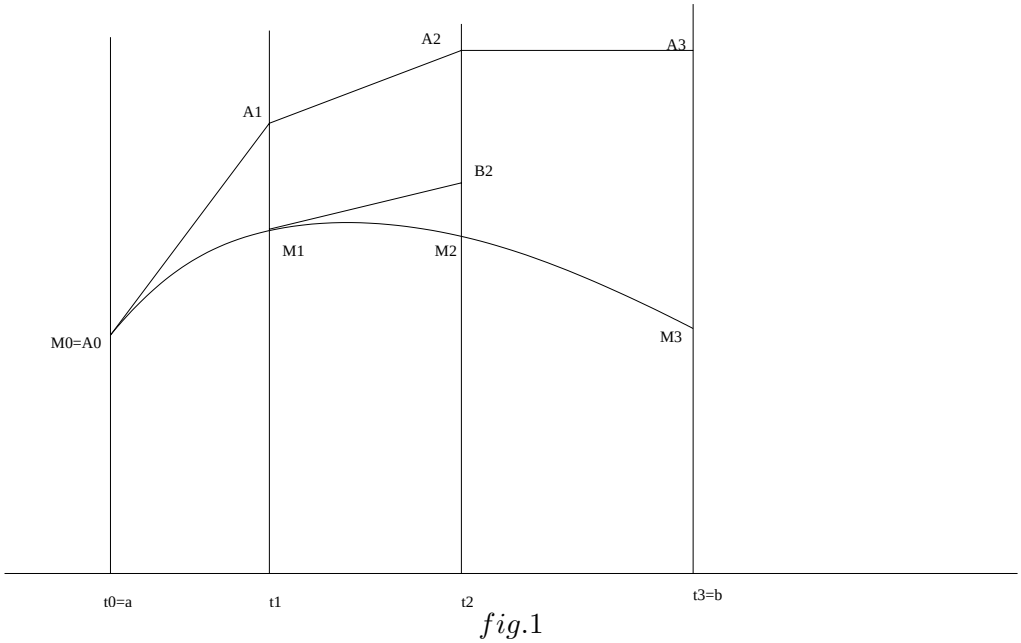

*fig.*1

Noter que la pente $f(t_n, y_n)$ du segment $A_n A_{n+1}$ diffère en général (sauf pour $n = 0$) de $z'(t_n) = f(t_n, z(t_n))$. Par exemple sur la figure $z'(t_1)$ est la pente du segment $[M_1 B_2]$, et $f(t_1, y_1)$ celle du segment $[A_1 A_2]$

Exercice : Montrer que dans la situation du théorème 1, toute solution approchée affine par morceaux construite par la méthode de la tangente a un graphe contenu dans le cylindre de sécurité C.

Nous admettrons le résultat suivant de convergence uniforme, qui implique le théorème 1.

Soit $y_p(t)$ une suite de solution approchées, construite par la méthode d'Euler à partir d'une subdivision $t_{p,0} < t_{p,1} < \cdots < t_{p,N_p}$ de l'intervalle $[t_0, t_0 + T]$. On note $h_{p,max} = max(h_{p,1}, \cdots, h_{p,N_p})$ appelé pas maximal de la subdivision pour y_p.

Dans la situation du théorème 1 *toute suite* $y_p(t)$, de solutions approchées, dont le pas maximal tend vers zéro lorsque $p \to \infty$, converge uniformément vers *la* solution $z(t)$.

Dans le chapitre suivant on se place uniquement dans la situation du théorème 1, avec unicité de $z(t)$.

2 Introduction à la notion de schéma numérique

On considère une équation différentielle $y' = f(t, y)$, dans laquelle f satisfait aux conditions du théorème de Cauchy-Lipschitz, et on notera dans toute la suite $t \mapsto z(t)$ la solution unique sur $[t_0, t_0 + T]$, dont le graphe est contenu dans un cylindre de sécutité $[t_0 - T, t_0 + T] \times \overline{\mathcal{B}}(y_0, r_0)$.

On appelera parfois z la "solution exacte", par opposition aux solutions approchées notées $t \mapsto y(t)$. Pour l'évaluation approchée de $z(t)$, la stratégie générale est la suivante :

- Définir un schéma numérique : on appelle ainsi les formules de récurrence associées à une méthode numérique de résolution des équations différentielles. Il s'agit d'une méthode de calcul de valeurs approchées notées y_n des $z(t_n)$, où les t_n sont les termes d'une subdivision à N pas de $[t_0, t_0+T]$:

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

  Le $n^{ième}$ pas est noté $h_n = t_{n+1} - t_n$ (on commence à l'indice 0), et le pas maximum est $h_{\max} = \max_{0 \leq n \leq N-1} h_n$
  On peut considérer la fonction affine par morceaux $y(t)$ telle que pour tout n, $y(t_n) = y_n$, ce qui donne précisément :

  $$y(t) = y_n + \frac{t - t_n}{h_n}(y_{n+1} - y_n) \quad \text{si} \quad t \in [t_n, t_{n+1}]$$

- Evaluer $\|y(t) - z(t)\|$ pour $t \in [t_n, t_{n+1}]$ et trouver un majorant (en fonction de f et de la méthode utilisée) de

  $$\sup_n \left(\sup_{t \in [t_n, t_{n+1}]} \|y(t) - z(t)\| \right)$$

- Etudier le comportement de cette majoration en fonction de la subdivision et particulièrement lorsque $h_{\max} \to 0$.

Définition 1 *Un schéma numérique à un pas explicite est une équation de récurrence de la forme :*

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

Le domaine de définition de Φ contient au moins $U \times \{0\}$ et on doit vérifier dans chaque situation concrète que l'itération est compatible avec ce domaine de définition.

Mentionnons d'autres types de schémas numériques dont l'étude dépasse le cadre de ce cours. Dans les *méthodes implicites* la fonction Φ dépend aussi de y_{n+1}. Dans un *schéma numérique explicite à q pas* Φ dépend aussi d'un nombre fixe de termes précédemment calculés, $y_{n-q+1} \dots y_n$, et la méthode doit être complétée par une *initialisation*, pour le calcul des q premiers termes.

Définition 2 *Un schéma numérique à un pas implicite est de la forme :*

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

Dans le cas d'une méthode implicite 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 solution unique du moins pour tout h assez petit. Dans les cas les plus courants cela résultera du théorème des fonctions implicites.

Définition 3 *Un schéma numérique à q pas explicite est de la forme :*

$$\begin{cases} y_{n+1} = y_n + h_n \Phi(t_n, y_n, \cdots, y_{n-q}, h_n) \\ t_{n+1} = t_n + h_n \end{cases}$$

avec $n = Nq, \cdots, N-1$.

Bien sur la calcul de y_n n'est possible qu'à partir de l'indice q et la méthode doit être complétée par une *initialisation*, le calcul des q premiers termes, par exemple par une méthode à un pas.

L'erreur de consistance au pas n est par définition l'erreur commise sur y_{n+1}, lorsqu'on prend pour les valeurs précédentes des y_k les valeurs exactes $z(t_k)$, ce qui donne la définition suivante où nous n'explicitons que pour les méthodes a un pas.

Définition 4 L'erreur de consistance *est la suite*

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

Illustration sur la figure 1. Par définition les erreurs de consistance e_1, e_2 *etc*, sont des différences d'ordonnées égales respectivement aux mesures des segments orientés $\overline{M_1 A_1}$, $\overline{M_2 B_2} \cdots$. Au premier pas e_1 n'est autre que l'erreur $y_1 - z(t_1)$, mais dès le deuxième pas l'erreur $y_2 - z(t_2) = \overline{M_2 A_2})$ diffère bien sur de e_2 car $A_2 \neq B_2$, mais aussi du cumul $e_1 + e_2$ dès que les pentes $f(t_1, y_1)$, et $z'(t_1) = f(t_1, z(t_1))$ des segments $A_1 A_2$ et $M_1 B_2$ diffèrent.

Les valeurs de $z(t_i)$ n'étant pas connues pour $i \geq 1$, l'erreur de consistance est surtout un outil théorique qui intervient dans le calcul d'une majoration de l'erreur $|z(t_n) - y_n|$.

Définition 5 *On dit qu'un schéma numérique est d'ordre p, si $e_n = o(h_n^{p+1})$, lorsque $h_n \to 0$.*

Comme on le verra dans la section 4.1.3, L'ordre d'une méthode est une indication importante qui avec la propriété dite de stabilité qui dépend de f gouverne la convergence de $y(t)$ vers $z(t)$. Nous conclurons seulement cette section par un résultat admis qui donne une première idée intuitive de la notion d'erreur de consistance :

Définition 6 *Une méthode numérique est dite consistante si*

$$\lim_{h_{\max} \to 0} \sum_{n=0}^{N-1} |e_n| = 0$$

Théorème 2 *.- Une méthode à un pas est consistante si et seulement si quel que soit $(t, y) \in U$, on a :*

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

Démonstration.- Nous donnons plus loin une démonstration détaillée qui fait appel à un maniement de sommes de Riemann. Intuitivement on peut justifier cet enoncé en remarquant que par le théorème des accroissements finis, $z(t_{n+1}) - z(t_n)$ est de l'ordre de $h_n z'(t_n) = h_n f(t_n, z(t_n))$, donc que l'erreur de consistance $e_n = z(t_{n+1}) - z(t_n) - h_n \Phi(t_n, z(t_n), h_n)$ est de l'ordre de $\Big[f(t_n, z(t_n)) - \Phi(t_n, z(t_n), h_n) \Big] . h_n$. Si la condition de l'énoncé n'est pas remplie cette différence est pour h tendant vers zéro de l'ordre de h, avec un facteur multiplicatif ne tendant pas vers zéro. Le cumul des erreurs d d'arrondi serait donc de l'ordre du cumul des h_n c'est à dire de la quantité fixe T, donc l'erreur ne tendrait pas vers 0 avec le pas. □

3 Quelques exemples de méthodes à un pas

3.1 Méthode d'Euler.

On a déjà décrit cette méthode appelée aussi méthode de la tangente qui correspond au cas de la fonction :

$$\phi(t, y, h) = f(t, y)$$

indépendante de t, définie sur $U \times \mathbb{R}$.

Calcul de l'erreur de consistance

$$\begin{aligned} e_n =& z(t_n + h_n) - z(t_n) - h_n.f(t_n, z(t_n)) \\ =& z(t_n + h_n) - z(t_n) - h_n.z'(t_n) \qquad \text{par définition de } z \end{aligned}$$

Lorsque f est de classe $\mathcal{C}^1$, z est de classe $\mathcal{C}^2$ et l'erreur de consistance prend la forme suivante grâce à la formule de Taylor.

$$\begin{aligned} e_n =& \frac{1}{2}h_n^2 z''(t_n) + o(h_n^2) \\ & \frac{1}{2}h_n^2 f^{[1]}(t_n, z(t_n)) + o(h_n^2) \end{aligned}$$

Ainsi la méthode d'Euler est d'ordre un.

3.2 Méthode de Taylor d'ordre p.

On suppose ici que f est de classe $\mathcal{C}^p$. On a vu alors que z est de classe $\mathcal{C}^{p+1}$ et on a défini des fonction $f^{[k]}$, construite par récurrence à partir de f et de ses dérivées partielles telles que $z^{(k)}(t) = f^{[k-1]}(t, z(t)),$ pour $k = 1, \ldots, p+1$. La formule de Taylor à l'ordre $p+1$ s'écrit alors :

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

ou avec la formule de Taylor Lagrange :

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

Ceci suggère le schéma numérique suivant obtenu en remplaçant les valeurs inconnues $z(t_k)$ par les y_k.

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

La fonction Φ associée à cette méthode est :

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

Le résultat suivant généralise celui qu'on a déjà trouvé pour la méthode d'Euler $(\mathcal{T}_1)$.

Proposition 3 *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 une 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 de Taylor avec reste de Lagrange) :

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

3.3 Méthode du point milieu.

Notons M_n le point de coordonnées $(t_n, z(t_n))$ du graphe de z. Le segment $[M_n, M_{n+1}]$ a une pente plus proche en général de $z'(t_n + \frac{h_n}{2})$ pente de la tangente au "point milieu" que de $z'(t_n)$, pente de la tangente en M_n, comme le montre la figure ci-dessous :

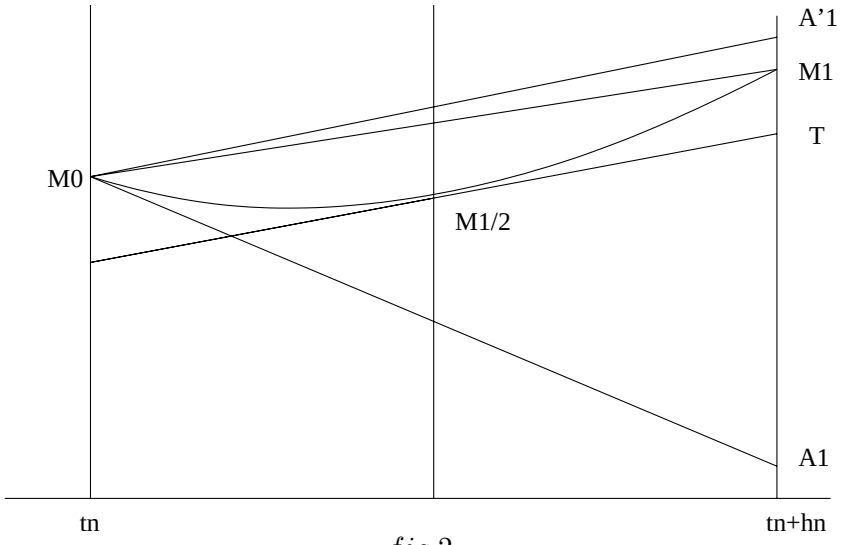

fig.2

On peut donc considérer qu'une approximation de $z(t_{n+1})$ à partir de $z(t_n)$ meilleure que l'expression $z(t_n) + f(t_n, z(t_n))$ de la méthode d'Euler est :

$$(1) \quad z(t_n) + h_n z'(t_n + \frac{h_n}{2}) = z(t_n) + h_n f(t_n + \frac{h_n}{2}, z(t_n + \frac{h_n}{2}))$$

Dans le schéma numérique que nous avons en vue on peut prendre par récurrence y_n approximation de $z(t_n)$.

Comme $z(t_n + \frac{t_n}{2})$ n'est pas connu non plus, il convient d'en chercher une approximation notée $y_{n+\frac{1}{2}}$. Le schéma d'Euler suggère de prendre $y_{n+\frac{1}{2}} = y_n + \frac{h_n}{2} f(t_n, y_n)$.

On aboutit ainsi donc au schéma numérique :

$$(\mathcal{M}) \quad \begin{cases} y_{n+\frac{1}{2}} = y_n + \dfrac{h_n}{2} f(t_n, y_n) \\ p_n = f(t_n + \dfrac{h_n}{2}, y_{n+\frac{1}{2}}) \\ y_{n+1} = y_n + h_n p_n \\ t_{n+1} = t_n + h_n \end{cases}$$

Ce schéma est encore une méthode à un pas explicite dans laquelle l'expression explicite de Φ obtenue en développant p_n est :

$$\Phi(t, y, h) = f(t_n + \frac{h_n}{2}, y_n + \frac{h_n}{2} f(t_n, y_n))$$

Proposition 4 *La méthode* $(\mathcal{M})$ *est d'ordre* 2 *dès que* f *est de classe* $\mathcal{C}^2$.

Démonstration.- On écrit le schéma $(\mathcal{M})$ avec $y_n = z(t_n)$ ce qui donne l'erreur de consistance

$$e_n = z(t_n + h_n) - z(t_n) - h_n p_n, \text{ avec } p_n = f(t_n + \frac{h_n}{2}, y_n + \frac{h_n}{2} f(t_n, z(t_n))$$

Comme p_n est une approximation de $z'(t_n + \frac{h_n}{2})$, on a interêt à décomposer e_n ainsi :

$$e_n = e'_n + e''_n \text{ avec } \begin{cases} e'_n = z(t_n + h_n) - z(t_n) - h_n z'(t_n + \frac{h_n}{2}) \\ e''_n = h_n(z'(t_n + \frac{h_n}{2}) - p_n). \end{cases}$$

On trouve d'abord en appliquant la formule de Taylor jusqu'au terme en h_n^3 aux deux fonctions z et z' :

$$\begin{aligned} e'_n &= z(t_n + h_n) - z(t_n) - h_n z'(t_n) - h_n(z'(t_n + \frac{h_n}{2}) - z'(t_n)) \\ &= \frac{h_n^2}{2} z''(t_n) + \frac{h_n^3}{6} z^{(3)}(t_n) + o(h_n^3) - h_n(\frac{h_n}{2} z''(t_n) + \frac{h_n^2}{8} z^{(3)}(t_n) o(h_n^2)) \\ &= \frac{h_n^3}{24} z^{(3)}(t_n) + o(h_n^3) = \frac{h_n^3}{24} f^{[2]}(t_n, z(t_n)) + o(h_n^3) \end{aligned}$$

d'autre part on peut appliquer la formule de Taylor par rapport à la variable y dans l'expression de e''_n, ce qui donne :

$$\begin{aligned} e''_n &= h_n[f(t_n + \frac{h_n}{2}, z(t_n + \frac{h_n}{2})) - f(t_n + \frac{h_n}{2}, z(t_n) + \frac{h_n}{2} f(t_n, y_n))] \\ &= h_n \frac{\partial f}{\partial y}(t_n + \frac{h_n}{2}, \Theta_n)[z(t_n + \frac{h_n}{2})) - z(t_n) - \frac{h_n}{2} f(t_n, y_n)] \end{aligned}$$

avec Θ_n sur le segment $]z(t_n) + \frac{h_n}{2} f(t_n, y_n), z(t_n + \frac{h_n}{2}))[=]y_{n+\frac{1}{2}}, z(t_n + \frac{h_n}{2}))[$ Du fait que $\Theta_n = y_n + O(h_n)$, et du fait que $\frac{\partial f}{\partial y}$ est de classe $\mathcal{C}^1$ on déduit que $\frac{\partial f}{\partial y}(t_n + \frac{h_n}{2}, \Theta_n) = \frac{\partial f}{\partial y}(t_n, y_n) + O(h_n)$ et finalement tenant compte de $z'(t_n) = f(t_n, z(t_n))$, et en appliquant encore la formule de Taylor à l'ordre 2 à $z(t_n + \frac{h_n}{2})$:

$$\begin{aligned} e''_n &= h_n(\frac{\partial f}{\partial y}(t_n, y_n) + O(h_n))(\frac{h_n}{2} z'(t_n) + \frac{1}{2}\frac{h_n^2}{2} z''(t_n) - \frac{h_n}{2} f(t_n, y_n)) \\ &= \frac{h_n^3}{8} \frac{\partial f}{\partial y}(t_n, y_n) z''(t_n) o(h_n^3) = \frac{h_n^3}{8} \frac{\partial f}{\partial y}(t_n, y_n) f^{[1]}(t_n, z(t_n)) + o(h_n^3) \end{aligned}$$

□ En apparence ce calcul semble restreint au cas scalaire $m = 1$, mais en fait il est intégralement valable en général à condition de considérer que $\frac{\partial f}{\partial y}(t, y).v$ désigne la valeur sur le vecteur $v \in \mathbb{R}^n$ de l'application linéaire $\mathbb{R}^n \to \mathbb{R}^n$ dérivé partielle de f au point $(t, y) \in U \subset \mathbb{R} \times \mathbb{R}^n$.
Exercice Montrer avec les même calculs mais la formule de Taylor-Lagrange la majoration explicite suivante, qui met en jeu des normes du sup sur le cylindre de sécurité :

$$e_n \leq (\frac{1}{24} \|f^{[2]}\|_\infty + \frac{1}{8} \|\frac{\partial f}{\partial y}\|_\infty \|f^{[1]}\|_\infty) h_{\max}.$$

Un exemple simple : On considère l'équation $y' = -y$, avec les données initiales $t_0 = 0,\ y_0 = 1$, et la subdivision de $[0, 1]$ de pas constant $\frac{1}{10}$, $t_n = \frac{n}{10},\ n = 0, \cdots, 10$.

La solution exacte est connue : $z(t) = e^{-t}$, ce qui permet de comparer aisément différentes méthodes quant à leur précision en fonction du pas.

La comparaison est proposée en exercice et on constatera des erreurs respectives de l'ordre de 2.10^{-2}, 10^{-3}, 2.10^{-5} pour les schémas $(\mathcal{T}_1)$ $(\mathcal{T}_2)$ et $(\mathcal{T}_3)$. On verra aussi que pour atteindre la même précision que par 10 pas avec $(\mathcal{T}_3)$, le nombre de pas nécessaires serait de :
-100 pas dans le cas de la méthode de Taylor d'ordre 2.
-10000 pas dans le cas de la méthode d'Euler.

Vérifier au passage que la méthode du point milieu donne *sur cet exemple* la même formule que $\mathcal{T}_2$

Cet exemple montre que la méthode d'Euler n'est pas assez précise pour qu'on puisse s'en contenter dans la pratique. L'augmentation du nombre de pas est en effet préjudiciable à cause du cumul des erreurs d'arrondis, qui pour une méthode donnée impose une borne à la précision.

Le choix d'une méthode résulte d'un compromis entre la sophistication du schéma utilisé qui doit rester raisonnable et le gain d'efficacité.

On étudiera dans la section 6 le méthode(s) de Runge Kutta. Il y a toute une famille de schémas numériques qui portent ce nom, mais le plus classique et le plus utilisé d'entre eux est celui d'ordre 4.

4 Etude générale de la convergence des méthodes à un pas : consistance et stabilité

4.1 Consistance, stabilité et convergence

4.1.1 La notion de consistance d'un schéma

Il s'agit du bon comportement de l'erreur du même nom.

Définition 7 *Une méthode numérique est dite consistante si*

$$\lim_{h_{\max}\to 0} \sum_{n=0}^{N-1} |e_n| = 0$$

4.1.2 La notion de stabilité

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

1) *Erreurs d'arrondi* : on représente en machine la valeur y_n issue du calcul par un nombre décimal à q chiffres : $|\widetilde{y_n} - y_n|$, est alors l'erreur d'arrondi majoré en valeur relative par $10^{-q} y_n$.

2) *Incertitude expérimentale* Dans la plupart des problèmes concrets la "vraie" valeur (notion mythique...) de y_0 est remplacée par une valeur $\widetilde{y_0}$ tirée d'une expérience, d'une hypothèse etc : $|\widetilde{y_0} - y_0|$ est donc majorée par un nombre qui dépend de la précision expérimentale.

La méthode ne peut donc être utile que si la perturbation sur $|\widetilde{y_N} - y_N|$ provoquée par une faible perturbation $|\widetilde{y_0} - y_0|$ des données initiales et par les erreurs d'arrondi sur les termes $\widetilde{y_n}$ antérieurs est faible :

Définition 8 *Une méthode est dite* stable *s'il existe $S > 0$ tel que quelle que soient les suites , définies par récurence par les formules :*

$$\begin{cases} y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n) \\ \widetilde{y_{n+1}} = \widetilde{y_n} + h_n \Phi(t_n, \widetilde{y_n}, h_n) + \epsilon_n, \end{cases}$$

on a :

$$\max_{0\le n\le N} |\widetilde{y_n} - y_n| \le S(|\widetilde{y_0} - y_0| + \sum_{n=0}^{N-1} |\epsilon_n|).$$

Remarques sur l'évaluation de l'erreur. Dans ces formules ϵ_n est une erreur d'arrondi *qu'on majore dans la pratique* en erreur relative par $0,5.10^{-q}$ en fonction de la précision de la machine : q est le nombre de chiffres dans l'écriture en virgule flottante et si $y_n \in [10^{k-1}, 10^k[$ avec $k \in \mathbb{Z}$, l'erreur d'arrondi absolue est au plus $0,5.10^{k-q}$. La quantité $|\tilde{y_0} - y_0|$, de son coté doit être évaluée (donc majorée) en fonction de la nature (physique ou autre) du problème modélisé.

Ainsi, selon la définition, on ne peut donc pas espérer une précision relative à priori meilleure que :

$$S \times (N+1) \times 0,5.10^{-16}$$

pour une écriture de réels avec 16 chiffres significatifs.

Par conséquent, le fait qu'une méthode soit stable n'est pas une garantie d'obtenir des résultats numériquement fiables lorsqu'on se heurte à l'un des deux écueuils suivants :
- Lorsqu'on on prend un pas très petit, donc N très grand N le cumul des erreurs d'arrondis peut provoquer une erreur trop élevée.
- La constante de stabilité S peut être très grande de l'ordre de 10^{+16} ou plus ce qui ote toute crédibilité aux résultats obtenus. C'est le cas de problème dits *numériquement mal posés*. C'est aussi le cas, lorsqu'on cherche les solutions pour $[t_0, t_0 + T]$, avec T trop grand. En effet que S croît exponentiellement avec T. Voir la formule finale dans la démonstration du théorème 5 avec une constante de stabilité en $S = e^{LT}$, et la sous section 4.3.2 avec le facteur $SCT = e^{LT}CT$ pour le facteur d'amplification de l'erreur.

4.1.3 La notion de convergence

Faisant abstraction des contraintes pratiques que nous venons d'évoquer on a quand même les résultats théoriques suivants, de démonstration très facile, qui relient les deux notions de consistance et de stabilité à la convergence de la méthode :

Définition 9 *Une méthode numérique est dite* convergente *si pour toute solution exacte z 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 :*

$$\lim_{\substack{h_{\max} \to 0 \\ y_0 \to z(t_0)}} \max_{0 \le n \le N} |y_n - z(t_n)| = 0.$$

Théorème 3 *Une méthode numérique à un pas qui est stable et consistante est convergente.*

Démonstration.- Posons $\widetilde{y_n} = z(t_n)$. Dans ce cas l'erreur de consistance est par définition le réel e_n qui complète la formule :

$$z(t_{n+1}) = \widetilde{y_{n+1}} = \tilde{y_n} + h_n \Phi(t_n, \tilde{y_n}, h_n) + e_n$$

Donc e_n joue pour la suite des $z(t_n)$ le role de la correction ϵ_n associé en général à une suite $\widetilde{y_n}$. D'après la définition de la consistance on a donc :

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

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

Exercice : Déduire du théorème 4.1.3 la convergence uniforme de $y(t)$, fonction linéaire par morceaux telle que $y(t_n) = y_n$, vers $z(t)$ en utilisant la continuité uniforme de z sur $[t_0, t_0 + T]$.

4.2 Quelques conditions suffisantes ou nécessaires et suffisantes de consistance ou de stabilité.

Dans cet énoncé on suppose que Φ remplit la condition suivante presque toujours réalisée dans les exemples usuels :

$$\Phi \text{ est continue sur un ouvert contenant } U \times [-h_0, h_0]$$

Théorème 4 .- *Une méthode à un pas est consistante si et seulement si quel que soit $(t, y) \in U$, on a :*

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

Démonstration.- D'après le théorème des accroissement finis, il existe quel que soit n un réel $c_n \in]t_n, t_{n+1}[$ tel que :

$$\begin{aligned} e_n &= z(t_{n+1}) - z(t_n) - h_n\Phi(t_n, z(t_n), h_n) \\ &= h_n z'(t_n) - h_n\Phi(t_n, z(t_n), h_n) = h_n[f(c_n, z(c_n)) - \Phi(t_n, z(t_n), h_n)] \\ &h_n[f(c_n, z(c_n)) - \Phi(c_n, z(c_n), 0)] + h_n[\Phi(c_n, z(c_n), 0) - \Phi(t_n, z(t_n), h_n)] \end{aligned}$$

Pour clarifier les calculs qui suivent on écrira au moment opportun ce résultat sous la forme :

$$\begin{aligned} e_n &= h_n(A_n + B_n) \\ A_n &= f(c_n, z(c_n)) - \Phi(c_n, z(c_n), 0) \\ B_n &= \Phi(c_n, z(c_n), 0) - \Phi(t_n, z(t_n), h_n) \end{aligned}$$

D'après l'uniforme continuité de Φ sur $C \times [-h_0, h_0]$ où C est le polycylindre de sécurité sur lequel on travaille :

$$\forall \epsilon > 0,\ \exists \eta > 0,\ \alpha > 0,\ \text{ tels que } (|h| < \eta,\ |t - t'| < \eta,\ |y - y'| < \alpha \ \Rightarrow\ |\Phi(t, y, 0) - \Phi(t', y', h)|)$$

Par ailleurs, quitte à diminuer η, on peut s'assurer en utilisant l'uniforme continuité de z sur $[t_0, t_0 + T]$ que $|h_n| < \eta \ \Rightarrow\ |z(c_n) - z(t_n)| < \alpha$, donc finalement en enchainant les deux implications précédentes par transitivité :

$$|h_n| < \eta \ \Rightarrow\ |\Phi(c_n, z(c_n), 0) - \Phi(t_n, z(t_n), h_n)| < \epsilon$$

On en tire

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

Autrement dit avec les notations abrégées $e_n = h_n(A_n + B_n)$, on a obtenu

$$h_{\max} < \eta \ \Rightarrow \sum_{n=0}^{N-1} h_n|B_n| < \epsilon T$$

Donc $\sum_{n=0}^{N-1} h_n|B_n|$ tend vers zéro quand $h_{\max} \to 0$. Or on peut par ailleur écrire les majorations suivantes :

$$\begin{aligned} |\sum_{n=0}^{N}(|e_n| - h_n|A_n|)| &\le \sum_{n=0}^{N} ||e_n| - h_n|A_n|| \\ &\le \sum_{n=0}^{N} |e_n - h_n.A_n| = \sum_{n=0}^{N} h_n|B_n| \end{aligned}$$

On a donc démontré que $\sum_{n=0}^{N}|e_n| - \sum_{n=0}^{N} h_n|A_n|$ tend vers zéro avec $h_{\max}$. Or on reconnait dans la suite

$$\sum_{n=0}^{N} h_n|A_n| = \sum_{n=0}^{N} h_n|f(c_n, z(c_n)) - \Phi(c_n, z(c_n), 0)|$$

une suite de sommes de Riemann de l'intégrale $I = \int_{t_0}^{t_0+T} |f(t, z(t)) - \Phi(t, z(t), 0)|$, donc une suite qui tend vers I. Le résultat obtenu nous donne alors aussi :

$$\lim_{h_{\max}} \sum_{n=0}^{N} |e_n| = I$$

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

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

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

Théorème 5 .- *Une condition suffisante de stabilité :*

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

Démonstration.- On reprend les notations de la définition 8 et on pose :

$$\theta_n = |\widetilde{y_n} - y_n|$$

Par définition de la constante de Lipschitz pour Φ.

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

quel que soient (t, y_1, h) et (t, y_2, h) dans le domaine de définition de Φ. La majoration suivante découle aussitôt de la définition de la suite de θ_n et de la condition de Lipschitz :

$$\theta_{n+1} = |\widetilde{y_{n+1}} - y_{n+1}| = |\tilde{y_n} - y_n + h_n(\Phi(t_n, \tilde{y_n}, h_n) - \Phi(t_n, y_n, h_n)) + \epsilon_n| \quad (1)$$

$$\leq (1 + h_n L)\theta_n + |\epsilon_n| \quad (2)$$

Lemme 1 *: Lemme de Gronwall discret .- Les inégalités (2) impliquent :*

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

Démonstration.- C'est un exercice élementaire sur les fonctions $\mathbb{R} \to \mathbb{R}$ de vérifier que $\forall x \in \mathbb{R}$, $1 + x \leq e^x$ et on a donc

$$1 + h_n L \leq e^{Lh_n}$$

Le lemme se déduit par une récurrence sans mystère de cette majoration et de l'inégalité (2).

L'étape de récurrence de n à $n+1$ s'écrit ainsi :

$$\begin{aligned}
\theta_{n+1} &\leq (1+h_nL)\theta_n + |\epsilon_n| \\
&\leq (1+h_nL)\Big[e^{L(t_n-t_0)}\theta_0 + \sum_{i=0}^{n-1} e^{L(t_n-t_{i+1})}|\epsilon_i|\Big] + |\epsilon_n| \\
&\leq e^{Lh_n}\Big[e^{L(t_n-t_0)}\theta_0 + \sum_{i=0}^{n-1} e^{L(t_n-t_{i+1})}|\epsilon_i|\Big] + |\epsilon_n| \\
&= e^{L(t_{n+1}-t_0)}\theta_0 + \sum_{i=0}^{n-1} e^{L(t_{n+1}-t_{i+1})}|\epsilon_i| + |\epsilon_n| \\
&= e^{L(t_{n+1}-t_0)}\theta_0 + \sum_{i=0}^{n} e^{L(t_{n+1}-t_{i+1})}|\epsilon_i|
\end{aligned}$$

□

On déduit de ce lemme que pour tout n,

$$\theta_n \leq e^{LT}(\theta_0 + \sum_{i=0}^{n-1} .|\epsilon_i|)$$

Ceci termine la démonstration du théorème de stabilité avec la constante de stabilité :

$$S = e^{LT}$$

□

Remarque.- 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 Φ est Lipschitzienne de constante L.

Corollaire 6 *Si f est Lipschitzienne en y, les méthodes d'Euler et du milieu sont convergentes.*

Démonstration.- D'après le théorème 4.1.3 il suffit pour cela d'établir la consistance et la stabilité.

- La consistance est une conséquence directe du théorème 4, et de l'égalité $\Phi_{|h=0} = f$. C'est aussi valable pour la méthode de Taylor T_p. La stabilité se déduit du théorème 5 et du fait que Φ est Lipschitzienne : pour la méthode d'Euler, c'est immédiat car $f = \Phi$, et pour la méthode du milieu cela résulte des calculs suivants :

$$\phi = f(t + \frac{h}{2}, y + \frac{h}{2}f(t,y))$$

Donc si f est Lipshitzienne en y de constante de Lipschitz L on obtient :

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

D'où le caractère lipschitzien de Φ avec la constante de Lipschitz $\Lambda = L + \frac{\delta}{2}L^2$, si on se limite a des pas h assez petits c'est à dire assujettis à une condition : $0 < h \leq \delta$ □

Note : On montre aussi que lorsque f est de classe $\mathcal{C}^p$ la méthode de Taylor $\mathcal{T}_p$ est convergente.

4.3 Ordre d'une méthode à un pas et erreur globale.

4.3.1 .-

Définition 10 *Une méthode consistante est dite d'ordre p si pour tout compact K il existe $C \geq 0$, tel que pour toute solution $z(t)$, de graphe $\{(t, z(t))\}$ contenu dans K, l'erreur de consistance satisfait à la condition :*

$$|e_n| \leq C h_n^{p+1}$$

Mentionnons une caractérisation de l'ordre p qui s'applique à la méthode de Taylor de même indice et généralise le critère de consistance déjà énoncé :

Proposition 5 *.- Sous l'hypothèse que Φ est de classe $\mathcal{C}^p$, la méthode est d'ordre p si et seulement si les conditions suivantes sont remplies :*

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

Démonstration.- Rappelons d'abord que l'erreur de consistance est :

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

La démonstration est similaire (en plus compliqué) a celle de la caractérisation de la consistance, qui correspond au cas $p = 1$. En appliquent la formule de Taylor Lagrange on a l'existence de $c_n, d_n \in]t_n, t_{n+1}[$ tels que :

$$\begin{aligned}
z(t_{n+1}) - z(t_n) =& h_n z'(t_n) + \cdots + \frac{h_n^k}{k!} z^{(k)}(t_n) + \cdots + \frac{h_n^p}{p!} z^{(p)}(t_n) + \frac{h_n^{p+1}}{(p+1)!} z^{(p+1)}(c_n) \\
=& \sum_{k=1}^{p} \frac{h_n^k}{k!} f^{[k-1]}(t_n, z(t_n)) + \frac{h_n^{p+1}}{(p+1)!} f^{[p]}(c_n, z(c_n)) \\
\Phi(t_n, z(t_n), h_n) =& h_n \big[\Phi(t_n, z(t_n), 0) + \cdots + \frac{h_n^\ell}{\ell!} \frac{\partial^\ell \Phi}{\partial h^\ell}(t_n, z(t_n), 0) + \cdots + \frac{h_n^{p-1}}{(p-1)!} \frac{\partial^{p-1} \Phi}{\partial h^{p-1}}(t_n, z(t_n), 0) \\
& + \frac{h_n^p}{p!} \frac{\partial^p \Phi}{\partial h^p}(d_n, z(d_n), d_n) \big]
\end{aligned}$$

On en tire

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

Pour que cette expression soit un DL en h_n de la forme $o(h_n^p)$, il faut et il suffit que tous les termes $\frac{f^{[\ell]}(t_n, z(t_n))}{\ell+1} - \frac{\partial^\ell \Phi(t_n, z(t_n), 0)}{\partial h^\ell})$ soient nuls ce qui fournit la condition de l'énoncé puisque par tout point (t_0, y_0) passe une unique solution locale. Le reste fournit la majoration de l'énoncé avec la constante $C = \frac{1}{(p+1)!} \|f^{[p]}\|_K + \frac{1}{p!} \|\frac{\partial^p \Phi}{\partial h^p}|_{K \times [0,\delta]}$, où les normes utilisées sont les normes du sup $\| \bullet \|_\infty$ et les solution sont supposée confinés dan un ensemble compact K. □

N.B. Pour justifier l'existence des $f^{[\ell]}$ on remarque que grâce à l'hypothèse de consistance on a $f = \Phi_| h = 0$, qui est donc bien de classe $\mathcal{C}^p$. Le cas $\ell = 0$ de la conclusion de l'énoncéredonne la condion de consistance, avec un démonstration simplifi'ee par le fait qu'on suppose ici Φ de classe $\mathcal{C}^1$ et pas seulement continue.

4.3.2 La constante SCT de majoration de l'erreur globale

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

$$\sum e_n \leq C \sum h_n^{p+1} \leq C(\sum h_n) h_{max}^p = CTh_{max}^p$$

En effet la somme de tous les pas est $\sum(t_{n+1} - t_n) = T$ si on travaille sur l'intervalle $[t_0, t_0 + T]$. Par définition de la stabilité, on trouve alors : $Max|y_n - z(t_n)| \leq S(|y_0 - z(t_0)| + CTh_{max}^p)$. L'erreur est donc majorée dans le cas d'absence d'erreur sur la condition initiale par

$$SCTh_{max}^p$$

4.3.3 La constante SCT de majoration de l'erreur globale

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

$$\sum e_n \leq C \sum h_n^{p+1} \leq C(\sum h_n) h_{max}^p = CTh_{max}^p$$

En effet la somme de tous les pas est $\sum(t_{n+1} - t_n) = T$ si on travaille sur l'intervalle $[t_0, t_0 + T]$. Par définition de la stabilité, on trouve alors : $Max|y_n - z(t_n)| \leq S(|y_0 - z(t_0)| + CTh_{max^p})$. L'erreur est donc majorée dans le cas d'absence d'erreur sur la condition initiale par

$$SCTh_{max}^p$$

5 Quelques remarques sur les aspects numériques

5.1 Sur les inconvénients du cumul des erreursd'arrondi

On considère la suite des valeurs arrondies des y_n notées $\tilde{y_n}$. Soit ρ_n l'erreur d'arrondi dans le calcul de $\Phi(t_n, \tilde{y_n}, h_n)$, et σ_n l'erreur d'arrondi dans l'évaluqtion de $\tilde{y_{n+1}}$, ce qui conduit à la formule :

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

Alors si S est une constante de stabilité et si ρ et σ sont des bornes pour les erreurs d'arrondi, on trouve :

$$\max |\tilde{y_n} - y_n| \leq S(|\epsilon_0| + \sum |h_n \rho_n + \sigma_n|) \leq S(|\epsilon_0| + T\rho + N\sigma)$$

En combinant avec la majoration de la dernière sous-section on trouve :

$$\max |\tilde{y_n} - z(t_n) \leq S(|\epsilon_0| + T\rho + N\sigma + CTh^p)$$

En nous plaçant pour simplifier dans l'hypothèse de pas constants, donc avec $N = \frac{T}{h}$, ce dernier majorant devient :

$$E(h) = S(|\epsilon_0| + T\rho) + ST(\frac{\sigma}{h} + Ch^p)$$

La fonction $E(h)$ passe par un minimum pour une valeur optimale de h, qu'il est inutile (nefaste) de dépasser puisque $\lim_{h \to 0} E(h) = +\infty$. Cette valeur est $h_{opt} = (\frac{\sigma}{C})^{\frac{1}{p+1}}$. On trouve le majorant optimal : $E(h_{opt}) = \sigma^{\frac{p}{p+1}} C^{\frac{1}{p+1}} p^{\frac{1}{p+1}} (\frac{p+1}{p})$

5.2 Problèmes de Cauchy bien posés numériquement, probl'emes mal conditionnés

Exemple 0.- On a déjà vu que le problème de Cauchy :

$$y' = \sqrt{2|y|}, \;\; y(0) = 0$$

est mal posé mathématiquement car la solution n'est pas unique.

Dans les deux exemples suivants on se place dans les conditions du théorème de Cauchy, mais d'autres difficultés surgissent.

Exemple 1.-

$$\left\{ \begin{array}{lcl} y' & = & 3y - t \\ y_0 & = & \frac{1}{3} \end{array} \right. \quad \text{à } 10^{-n} \text{ près} \quad \text{calculer } y(10).$$

Dans cet exemple 10^{-n} représente la pr"ecision de la machine traité comme un majorant de l'erreur d'arrondi. On prendra donc les conditions initiales respectives :

$$y_0 = \frac{1}{3}, \;\; \text{et } \tilde{y_0} = \frac{1}{3} + \epsilon \;\; \text{avec } |\epsilon| < 10^{-n}$$

Les solutions exactes se calculent et valent : $z(t) = Ce^{3t} + t + \frac{1}{3}$, avec $C = z(0) - \frac{1}{3}$. On a donc à comparer les deux solutions

$$\left\{ \begin{array}{lcll} y(t) & = & t + \frac{1}{3} & \\ \tilde{y(t)} & = & t + \frac{1}{3} & +\epsilon e^{3t} \end{array} \right.$$

Donc $|y(t) - \tilde{y(t)}| = |y(0) - \tilde{y(0)}|e^{3t} = \epsilon e^{3t}$, ce qui donne la majoration :

$$|y(10) - \tilde{y(10)}| \leq 10^{-n} e^{30}$$

Conclusion : Le problème du calcul de $y(10)$ est mal posé numériquement si on travaille avec $n = 12$ chiffres significatifs, car le majorant $10^{-12}e^{30} \approx 10,7$ est excessif d l'oredre de grandeur de $y(10)$?

La notion d'être mal posé numériquement dépend bien sur de la longueur T de l'intervalle sur lequel on travaille, et aussi de l'exposant (L=3 ici) de l'exponentielle qui n'est autre que la constante de Lipschitz de f. Pout $T = 1$ (calcul de $y(1) = 4/3 + e^1$), le probolème serait bien posé avec une amplification de l'erreur d'arrondi initiale $e.10^{-12}$. De même le problème de $y(10)$ reste bien posé si on prend 16 chiffres significatifs et cesse de l'être à parti de $T = 13$ environ.

Toutes ces considération reflètent le fait que la constante de stabilité qu'on a trouvé dans le théorème 5 est en fonction de la constante de Lipschitz e^{LT}, donc la constante d'amplification trouvé dans la section précédente est $SCT = E^{LT}.C.T$

Exemple 2.-

$$\left\{ \begin{array}{lcl} y' & = & -150y + 30 \\ y_0 & = & \frac{1}{5} \end{array} \right. \quad \text{à } 10^{-n} \text{ près} \quad \text{calculer } y(10).$$

Cette fois le problème est bien posé numériquement car en conservant les notation analogues a celles de l'exemple 1 :

$$\tilde{y}(t) = \frac{1}{5} + \epsilon e^{-150t}$$

et le problème est bien posé numériquement puisque $|y(0) - \tilde{y(0)}|e^{-1500}$ est très petit même pour de grandes valeurs de ϵ. Ceci montre qu'en un certain sens le problème est trop bien posé : a l'inversse si on cherche a retrouver $\tilde{y_0}$, a partir d'une perturbation majorée α de $y(10) = \frac{1}{5}$ on trouve : +

$$tildey(10) - \frac{1}{5}| < \alpha \Rightarrow \;\; |\tilde{y(0)} - \frac{1}{5}| < \alpha.e^{+1500}$$

ce qui est numériquement impraticable.

On remarque que la constante de stabilité est a nouveau énorme : $S = e^{+150T} = e^{+1500}$, et $SCT = 10Ce^{+1500}$, et resterait excessive même pour des plus petites valeurs de T. On dit que ce problème est mal conditionné. Un tel problème même bien posé numériquement peut réserver des surprises à la mise en oeuvre de la métrhode d'Euler. On trouve :

$$y_{n+1} = -150y_n + 30, \text{ d'ou } |y_{n+1} - \frac{1}{5}| = -150|y_n - \frac{1}{5}|$$

donc $\tilde{y_n} = \frac{1}{5} + (1 - 150h)^n(\tilde{y_0} - \frac{1}{5})$. Pour $T = 1$, si on a l'imprudence de prendre un pas trop élevé la suite des $\tilde{y_n}$, s'ecarte radicalement de $\tilde{y}(t_n)$ très voisin de 1/5. Par exemple si $T = 1$, et $h = 1/50$, on trouve $\tilde{y_n} - \frac{1}{5} = (-2)^n(\tilde{y_0} - \frac{1}{5})$ d'ou pour $y(1)$, l'approximation $2^{50}(\tilde{y_0} - \frac{1}{5}) \approx 10^{15}(\tilde{y_0} - \frac{1}{5})$. Le choix de h est bien sur dicté par la convergence de la suite $(1 - 150h)^n$, soit : $0 < h < \frac{1}{75}$.

6 Méthode(s) de Runge Kutta.

6.1 description de la méthode

On considère comme d'habitude un problème de Cauchy

$$\left\{ \begin{array}{ccc} y' & = & f(t, y) \\ y(t_0) & = & y_0 \end{array} \right.$$

ave une solution exacte $z(t)$ sur $[T_0, t_0 + T]$ et une subdivision :

$$t_0 < \cdots t_N = t_0 + T$$

On part de l'expression intégrale de l'accroissement $z(t_{n+1}) - z(t_n)$, dans laquelle on ramène l'intervalle d'intégration de $[t_n, t_{n+1} = t_n + h_n]$, à $[0, 1]$, par le changement de variables $t = t_n + uh_n$.

$$\begin{aligned} z(t_{n+1}) - z(t_n) &= \int_{t_n}^{t_{n+1}} f(t, z(t))dt \\ &= \int_0^1 h_n f(t_n + uh_n, z(t_n + uh_n))du \end{aligned}$$

ou encore

$$z(t_{n+1}) - z(t_n) = h_n \int_0^1 g(u)du \tag{3}$$

avec $g(u) = f(t_n + uh_n, z(t_n + uh_n))$

L'idée est d'utiliser un opérateur d'intégration approché (O.I.A.) pour calculer l'intégrale qui apparaît dans l'équation (3) :

$$\int_0^1 g(u)du \sim \Sigma(g) = \sum_{i=1}^{q} b_i g(c_i)$$

Comme les valeurs des $g(c_i) = f(t_n + c_i h_n, z(t_n + c_i h_n)) = f(t_{n,i}, z(t_{n,i}))$ ne sont pas connues, il faut aussi évaluer la fonction z aux points $t_{n,i} = t_n + c_i h_n$ par un calcul similaire d'O.I.A. :

$$z(t_{n,i}) = z(t_n) + h_n \int_0^{c_i} g(u)du \tag{4}$$

On se contentera pour simplifier des méthodes explicites où l'O.I.A. choisi utilise les valeurs des $g(c_j) = f(t_{n,j}, z(t_{n,j}))$, $j = 1, \cdots, i - 1$ antérieurement calculées :

$$\int_0^{c_i} g(u)du \sim \Sigma_i(g) = \sum_{j=1}^{i-1} a_{i,j} g(c_i)$$

d'où les formules d'approximation :

$$\begin{cases} z(t_{n,i}) & \sim & z(t_n) + h_n \sum_{j=1}^{i-1} a_{i,j} g(c_i) \\ z(t_{n+1}) & \sim & z(t_n) + h_n \sum_{i=1}^{q} b_i g(c_i) \end{cases}$$

L'algorithme de Runge Kutta associé aux opérateurs d'intégration approché Σ et Σ_i s'obtient en remplaçant chaque $g(c_i)$ par des valeurs approchées notées $p_{n,j}$, puis $z(t_{n,i})$ par des valeurs approchées $y_{n,i}$, où on utilise l'O.I.A. Σ_i avec les $p_{n,j}$ au lieu des $g(c_j)$. Il reste à poser au $i$ème pas : $y_{n,i} = f(t_{n,i}, y_{n,i})$. Le passage de y_n à y_{n+1} se fait alors en utilisant de la même façon l'O.I.A. Σ.

On parle d'algorithme de type RK_q, pour indiquer le nombre des c_i. La méthode étant *explicite* $a_{i,j} = 0$ pour $j \geq i$ et on est donc dontraint à prendre $c_1 = 0$, et $p_{n,1} = f(t_n, y_n)$.(NB : RK_1 explicite n'est donc rien d'autre que la méthode d'Euler). Le détail de l'algorithme RK_q est donc le suivant :

- $c_1 = 0$, $t_{n,1} = t_n$, $y_{n,1} = y_n$, $p_{n,1} = f(t_n, y_n)$
- Pour $i = 2, \cdots, q$,
  $$\left[\begin{array}{rcl} t_{n,i} & = & t_n + c_i\, h_n \\ y_{n,i} & = & y_n + h_n \sum_{j=1}^{i-1} a_{i,j} p_{n,j} \\ p_{n,i} & = & f(t_{n,i}\ ,\ y_{n,i}) \end{array} \right.$$
- $$\left[\begin{array}{rcl} t_{n+1} & = & t_n + h_n \\ y_{n+1} & = & y_n + h_n \sum_{j=1}^{q} b_j p_{n,j} \end{array} \right.$$

Proposition 6 *Tout schéma de type RK_q définit une méthode à un pas explicite de la forme : $y_{n+1} = y_n + h_n \Phi(t_n, y_n, h_n)$.*

Il suffit en effet de vérifier par récurrence sur i l'existence de formules :

$$y_{n,i} = y_n + h_n \Phi_i(t_n, y_n, h_n),$$
$$p_{n,i} = Q_i(t_n, y_n, h_n).$$

On part de $\Phi_1 = 0$ et $Q_1 = f(t, y)$, et la récurrence se fait selon les formules :

$$\Phi_i = \sum_{j=1}^{i-1} a_{i,j} Q_j(t, y, h)$$
$$Q_i(t, y, h) = f(t + c_i h, y + h\Phi_i(t, y, h)).$$

et se conclut par $\Phi(t, y, h) = \sum_{j=1}^{q} b_j Q_j(t, y, h)$.

6.2 Quelques exemples

On fait systématiquement l'hypothèse que les O.I.A utilisés sont d'ordre au moins zéro(=exacts sur les fonctions constantes), ce qui setraduit par les égalités :

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

On présent usuellement les données sous forme d'un tableau :

$$\begin{array}{c|cccccc} c_1 = 0 & 0 & & & & & \\ c_2 & a_{2,1} & 0 & & & & \\ c_3 & a_{3,1} & a_{3,2} & 0 & & & \\ \vdots & & & & \ddots & & \\ \vdots & & & & & \ddots & \\ c_q & a_{q,1} & a_{q,2} & \cdots & \cdots & a_{q,q-1} & 0 \\ \hline 1 & b_1 & b_2 & \cdots & \cdots & b_{q-1} & b_q \end{array}$$

dans lequel la première colonne est la somme des suivantes.

– Les algorithmes de type RK_2 sont donc régis par un tableau du type suivant :

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

On pourra voir en TD que pour $\alpha \neq 0$ fixé la valeur optimale de b, pour laquelle l'ordre de la méthode est le plus grand est $b = 1 - \frac{1}{2\alpha}$.
Les O.I.A. de la méthode sont :

$$\int_0^\alpha g(u)du \sim \ \Sigma_2(g) = \alpha g(0) \tag{5}$$

$$\int_0^1 g(u)du \sim \ \Sigma(g)(1 - \frac{1}{2\alpha})g(0) + \frac{1}{2\alpha}g(1) \tag{6}$$

– Pour $\alpha = 1$, ce dernier O.I.A. est celui de la méthode des trapèzes et la méthode correspondante $(RK_2)_{\alpha=1}$ est connue sous le nom de *Méthode de Heun.*

– La méthode RK_4 classique correspond au tableau :

$$\begin{array}{c|cccc} 0 & 0 & & & \\ c_2 = \frac{1}{2} & \frac{1}{2} & 0 & & \\ c_3 = \frac{1}{2} & 0 & \frac{1}{2} & 0 & \\ c_4 = 1 & 0 & 0 & 1 & \\ \hline 1 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} \end{array}$$

Les O.I.A. de la méthode sont :

$$\int_0^{\frac{1}{2}} g(u)du \sim \ \Sigma_2(g) = \frac{1}{2}g(0)$$

$$\int_0^{\frac{1}{2}} g(u)du \sim \ \Sigma_3(g) = \frac{1}{2}g(\frac{1}{2})$$

$$\int_0^1 g(u)du \sim \ \Sigma_4(g) = g(\frac{1}{2})$$

$$\int_0^1 g(u)du \sim \ \Sigma(g) = \frac{1}{6}(g(0) + 2g(\frac{1}{2} + 2g(\frac{1}{2} + g(1))$$

Les coefficients sont ceux de l'O.I.A. de Simpson.
L'algorithme associé pour le calcul des y_n est le suivant :

$$\left[\begin{array}{rcl} t_{n,2} & = & t_n \ + \ \frac{1}{2}\, h_n \\ y_{n,2} & = & y_n + \frac{1}{2}h_n f(t_n, y_n) \\ p_{n,2} & = & f(t_{n,2}\ ,\ y_{n,2}) \end{array}\right] \quad \left[\begin{array}{rcl} t_{n,3} & = & t_n \ + \ \frac{1}{2}\, h_n \\ y_{n,3} & = & y_n + \frac{1}{2}h_n p_{n,2} \\ p_{n,3} & = & f(t_{n,3}\ ,\ y_{n,3}) \end{array}\right] \quad \left[\begin{array}{rcl} t_{n,4} & = & t_n \ + \ h_n \\ y_{n,4} & = & y_n + h_n p_{n,3} \\ p_{n,4} & = & f(t_{n,4}\ ,\ y_{n,4}) \end{array}\right]$$

$$\text{et} \ \ y_{n+1} = y_n + \frac{h_n}{6}(p_{n,1} + 2p_{n,2} + 2p_{n,3} + p_{n,4})$$

6.3 Thèmes de TD

6.3.1 Tester $(RK)_4$ sur l'exemple fétiche $y' = -y$.

Montrer qu'avec un pas $h = 0,1$, on trouve $y_n = (0,9048375)^n$, ce qui donne $y_{10} \simeq 0,3678798$, ce qui fournit la solution exacte étant e^{-t}, une approximation de $e^{-1} \simeq 0,3678794$ à mieux que 10^{-6} près.

Démonstration$L$ a reprise de l'algorithme de $(RK4)$ lorsque $f(t,y) = -y$ donne en effet ;

$$y_{n,2} = y_n - \frac{h_n}{2} y_n = -p_{n,2}$$

$$y_{n,3} = y_n + \frac{h_n}{2}(\frac{h_n}{2} - 1)y_n = y_n(1 - \frac{h_n}{2} + \frac{h_n^2}{4}) = -p_{n,3}$$

$$y_{n,4} = y_n - h_n(1 - \frac{h_n}{2} + \frac{h_n^2}{4})y_n = y_n(1 - h_n + \frac{h_n^2}{2} - \frac{h_n^3}{4}) = -p_{n,4}$$

et enfin

$$\begin{aligned} y_{n+1} =& y_n - \frac{h_n}{6} y_n(1 + 2(1 - \frac{h_n}{2}) + 2(1 - \frac{h_n}{2} + \frac{h_n^2}{4}) + 1 - h_n + \frac{h_n^2}{2} - \frac{h_n^3}{4}) \\ =& y_n - \frac{h_n}{6} y_n(6 - 3h_n + h_n^2 - \frac{h_n^3}{4}) \\ =& y_n(1 - h_n + \frac{h_n^2}{2} - \frac{h_n^3}{6} - \frac{h_n^4}{24}) \end{aligned}$$

Ainsi on trouve par récurrence $y_n = (1 - h_n + \frac{h_n^2}{2} - \frac{h_n^3}{6} - \frac{h_n^4}{24})^n$, ce qui en reportant $h = 0,1$, puis $n = 10$, fournit les valeurs numériques approchées indiquées. □
On remarque que sur cet exemple, le résultat est le même que celui que fournit la méthode de Taylor d'ordre 4.

6.3.2 Ordre d'une méthode $(RK)_4$.

La fonction $\Phi(t,y,h))$, s'obtient de la façon suivante : on réécrit l'algorithme de Runge Kutta en substituant au départ t et y et h à t_n et y_n et h_n, le même calcul fournissant alors des fonctions $y_i(t,y,h)$ et $p_i(t,y,h)$ au lieu des $y_{n,i}$ $p_{n,i}$ et la fonction cherchée provient de $y + h.\Phi(t,y,h))$ a la place de y_{n+1}.
– Pour $i = 2, \cdots, q$,

$$\left[\begin{array}{rcl} t_i & = & t \ + \ c_i \, h \\ y_i & = & y + h \sum_{j=1}^{i-1} a_{i,j} p_j(t,y,h) \\ p_i(t,y,h) & = & f(t_i \ , \ y_i(t,y,h)) \end{array}\right.$$

– $\Phi(t,y,h) = \sum_{j=1}^{q} b_j p_j(t,y,h)$
La méthode est consistante car on trouve aisément par récurrence sur i et pour tout i :

$$y_i(t,y,0) = y \qquad p_i(t,y,0) = f(t,y)$$

d'où $\Phi(t,y,0) = (\sum_{j=1}^{q} b_j) f(t,y) = f(t,y)$ ce qui est la condition suffisante trouvée dans le théorème 4.
Lemme 2

$$\frac{\partial \Phi}{\partial h}(t,y,0) = (\sum_{j=1}^{q} b_j c_j) f^{[1]}(t,y)$$

Démonstration$O$ n a $\Phi = \sum_{i=1}^{q} b_i f(t + c_i h, y_i(t,y,h))$ 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 $y_i(t,y,h) = y + \sum_{j=1}^{q} a_{i,j} f(t + c_j h, y_j(t,y,h))$, donc :

$$\frac{\partial y_i}{\partial h} = \sum_{j=1}^{q} a_{i,j} f(t + c_j h, y_j(t,y,h)) + h \sum_{j=1}^{q} a_{i,j} \frac{\partial}{\partial h}[f(t + c_j h, y_j(t,y,h))]$$

et puisque $y_i(t,y,0)=y$, on conclut par :

$$\frac{\partial y_i}{\partial h}(t,y,0)=(\sum_{j=1}^{q}a_{i,j})f(t,y)=c_i f(t,y).$$

et

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

□

On a vu que la condition nécessaire et suffisante pour que la méthode soit d'ordre 2 est : $\frac{\partial \Phi}{\partial h}(t,y,0)=\frac{1}{2}f^{[1]}(t,y)$. Au vu du lemme c'et équivalent à la condition numérique suivante :

$$\sum_{j=1}^{q}b_j c_j=\frac{1}{2}. \tag{7}$$

Exemple 1) On reprend l'exemple de (RK_2), avec le tableau :

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

La condition d'être d'ordre 2 est : $0\times b+\alpha(1-b)=\frac{1}{2}$, et on retrouve la valeur optimale $b=b_1=1-\frac{1}{2\alpha}$.

2) La méthode classique (RK_4) est d'ordre au moins deux également toujours d'après l'équation , puisque : $0\times\frac{1}{6}+\frac{1}{2}\frac{2}{6}+\frac{1}{2}\times\frac{2}{6}+1\times\frac{1}{6}=\frac{1}{2}$

On montre en fait que la méthode est d'ordre 4.

exercice Montrer en calculant $\frac{\partial^2\Phi}{\partial h^2}$, à comparer à $\frac{1}{3}f^{[2]}$, que la méthode (RK_4) est d'ordre ≥ 3.

On établira que en général les conditions pour qu'une méthode de Runge-Kutta d'ordre ≥ 2 soit en fait d'ordre ≥ 3 sont :

$$\sum_{j=1}^{q}b_j c_j^2=\frac{1}{3}\quad\text{et}\quad\sum_{i,j}b_i a_{i,j}c_j=\frac{1}{6}$$

On considère uine méthode de Runge Kutta de paramètres $a_{i,j}\ _{b_j}$, et on lui associe le coefficient réel positif :

$$\alpha=\max_i(\sum_j|a_{i,j}|)$$

Proposition 7 *Si $f(t,y)$ est Lipschitzienne de rapport k, la fonction Φ est Lipschitzienne de rapport : $\Lambda=k\sum_{j=1}^{q}|b_j|(1+\alpha hk+\cdots+(\alpha hk)^{j-1})$ et la méthode (RK_n) associée est stable.*

Démonstration On montre par récurrence sur i que la fonction $y_i(t,y,h)$ est Lipschitzienne de rapport

$$1+\alpha hk+\cdots+(\alpha hk)^{i-1}$$

En effet étant données deux condition initiales y et z la définition de y_i aboutit à :

$$\begin{aligned}|y_i(t,y,h)-y_i(t,z,h)| =&|y-z+h\sum_{j=1}^{q-1}a_{i,j}(f(t+c_j,y_j(t,y,h)-f(t+c_j,y_j(t,z,h))|\\ \leq&|y-z|+\sum_{j=1}^{q-1}|a_{i,j}|.|y_j(t,y,h)-y_j(t,z,h)|\\ \leq&|y-z|+(\alpha hk)\max_{j<i}|y_j(t,y,h)-y_j(t,z,h)|\end{aligned}$$

La conclusion pour la fonction y_i en découle. En passant à Φ on en déduit le résultat :

$$\begin{aligned}|\Phi(t,y,h)-\Phi(t,z,h)| =&|\sum_{j=1}^{q}b_j[f(t+c_j,y_j(t,y,h)-f(t+c_j,y_j(t,z,h)]|\\ \leq&\sum_{j=1}^{q}|b_j|.k|y_j(t,y,h)-y_j(t,z,h)|\\ \leq&|y-z|+(\alpha hk)\max_{j<i}|y_j(t,y,h)-y_j(t,z,h)|\leq k\Lambda|y-z|.\end{aligned}$$

□