4.2 Simulations numériques des EDO : schémas explicites

Nous ne considérons dans cette partie que le problème de Cauchy d'ordre 1 (4.15) ci-dessous (nous verrons en cours d'équations différentielles que quel que soit l'ordre d'une EDO, il est toujours possible de se ramener à l'ordre 1) suivant :

$$(\mathscr{C}) \quad \left\{ \begin{array}{rcl} x'(t) & = & f(t,x), \\ x(t_0) & = & x_0. \end{array} \right. \tag{4.15}$$

Mais attention, cette fois-ci $t \in [t_0, t_0 + T]$ (avec $T \in \mathbb{R}_+^*$), étant donné que l'on ne peut pas faire de simulation pendant un temps infini, le temps doit rester fini.
D'autre part, $x : [t_0, t_0 + T] \to \mathbb{R}^n$, $n \in \mathbb{N}$. Nous supposerons quasiment tout le temps pour les exercices d'applications que $n = 1$.

Remarque
Nous supposerons dans toute la suite, que la fonction $f : [t_0, t_0+T] \times \mathbb{R}^n \to \mathbb{R}^n$ est continue sur l'ensemble étudié, est globalement lipschitzienne par rapport à sa variable x sur l'intervalle étudié. En général, il nous suffira de montrer que la dérivée de f par rapport à sa variable x est bornée sur l'intervalle de t étudié. De telle sorte que nous n'aurons pas à nous soucier de l'existence globale et de l'unicité des solutions.

L'objectif est de calculer de façon approchée la solution du problème de Cauchy (4.15).
Pour cela nous discrétisons l'intervalle de temps $[t_0, t_0 + T]$ de la façon suivante : on définit une subdivision

$$t_0 < t_1 < t_2 < ... < t_N = t_0 + T,$$

et on cherche à définir

$$x_0, x_1, x_2, ..., x_N,$$

tels que les x_i, $i = 0, ..., N$, approchent les $x(t_i)$ (solution exacte aux points t_i).
On note

$$h_i = t_i - t_{i-1} > 0,$$

le i-ème pas de temps et $h = \max\limits_{i=1,...,N} h_i > 0$.

Remarque

1. *Dans toute la suite, nous supposerons les t_i régulièrement espacés. Autrement dit*

$$h = h_1 = h_2 = ... = h_N > 0,$$

2. *Si nous divisons l'intervalle $[t_0, t_0 + T]$ de longueur T, en N sous intervalles, nous aurons alors*

$$T = N.h \quad \textit{ou encore} \quad h = T/N \textit{ (avec } N \in \mathbb{N}^*).$$

Commençons par le schéma le plus simple : le schéma d'Euler progressif ou explicite.

4.2.1 Schéma d'Euler progressif (ou explicite)

On considère le problème de Cauchy ($\mathscr{C}$) (4.15). Il existe deux méthodes pour construire ce schéma.

Méthode 1

Pour un point $t = t_n$, quelconque avec $n = 0, 1, ..., N-1$, nous avons l'EDO qui satisfait

$$\begin{cases} x'(t_n) & = & f(t_n, x(t_n)), \\ x(t_0) & = & x_0 \text{ donné.} \end{cases}$$

On approche alors $f(t_n, x(t_n))$ par $f(t_n, x_n)$, et il faut approcher ensuite $x'(t_n)$. Pour cela il semble naturel d'approcher la dérivée par son taux d'accroissement aux valeurs approchées x_{n+1} et x_n, autrement dit

$$x'(t_n) \approx \frac{x(t_{n+1}) - x(t_n)}{t_{n+1} - t_n} \approx \frac{x_{n+1} - x_n}{t_{n+1} - t_n} = \frac{x_{n+1} - x_n}{h}.$$

Le schéma d'approximation du système ($\mathscr{C}$) au point t_n est donc

$$\begin{cases} \dfrac{x_{n+1} - x_n}{h} & = & f(t_n, x_n), \quad \text{pour } \ n = 0, ..., N-1, \\ x(t_0) & = & x_0. \end{cases}$$

soit encore,

$$(\mathscr{E}) \quad \begin{cases} x_{n+1} & = & x_n + hf(t_n, x_n), \quad \text{pour } \ n = 0, ..., N-1, \\ x(t_0) & = & x_0. \end{cases}$$

C'est ce que l'on appelle le schéma d'Euler progressif ou explicite. On dit qu'il est à un pas parce que x_{n+1} est défini seulement à partir du pas précédent x_n et non pas des autres pas précédents (on l'aurait alors appelé multipas).

Méthode 2

Une autre méthode pour construire ce schéma est d'utiliser la méthode des rectangles à gauche. En effet, rappelons le problème au point $t \in [t_0, t_0 + T]$:

$$\begin{cases} x'(t) & = & f(t, x(t)), \\ x(t_0) & = & x_0 \text{ donné.} \end{cases}$$

D'après le théorème de Cauchy-Lipschitz (et par les hypothèses que nous avons prises), nous avons x' et f continues. Nous pouvons donc intégrer l'équation différentielle entre t_n et t_{n+1}. Nous avons ainsi, pour tout $n = 0, 1, ..., N-1$,

$$\int_{t_n}^{t_{n+1}} x'(s)ds = \int_{t_n}^{t_{n+1}} f(s, x(s))ds.$$

L'intégrale du premier membre est connue, il s'agit du théorème fondamental du calcul, c'est la valeur $x(t_{n+1}) - x(t_n)$. L'intégrale du second membre est quant à elle approchée par la méthode des rectangles à gauches :

$$\int_{t_n}^{t_{n+1}} f(s,x(s))ds \approx f(t_n,x(t_n)).(t_{n+1}-t_n) = f(t_n,x(t_n)).h.$$

Nous avons alors

$$x(t_{n+1}) - x(t_n) \approx f(t_n,x(t_n)).h,$$

et comme nous approchons les $x(t_n)$ par x_n pour $n = 0, ..., N$, nous obtenons bien le même système qu'avec la première méthode, à savoir

$$(\mathscr{E}) \quad \left\{ \begin{array}{lcl} x_{n+1} & = & x_n + hf(t_n,x_n), \quad \text{pour} \;\; n = 0, ..., N-1, \\ x(t_0) & = & x_0. \end{array} \right.$$

L'avantage de cette méthode est qu'il est plus facile de comprendre la construction d'autres schémas que nous verrons dans la section 4.2.3 :

1. le schéma du point milieu où l'intégrale est approchée par le rectangle passant dont la hauteur est la solution exacte au milieu de l'intervalle,
2. le schéma de Heun donc l'approximation de l'intégrale se fait par la méthode des trapèzes.

Notons également que nous aurions pu prendre la méthode des rectangles à droite, nous n'aurions alors pas de schéma explicite mais un schéma d'Euler implicite (appelé aussi schéma rétrograde). Nous détaillerons ceci dans une des dernières sections de ce chapitre.
Nous voyons que les erreurs peuvent très vite s'accumuler. La question que l'on se pose est alors la suivantes :
est-ce que la solution approchée tend vers la solution de l'EDO quand h tend vers 0 (c'est à dire quand la discrétisation est très fine et les intervalles très petits) ?
Pour l'étudier nous allons introduire trois notions fondamentales dans ce chapitre : la consistance, la stabilité et la convergence pour des schémas explicites quelconques.

4.2.2 Schémas explicites à un pas

Définition 12 (SCHÉMAS EXPLICITES A UN PAS)

Un schéma explicite à un pas approchant la solution d'un problème de Cauchy $(\mathscr{C})$ donné par (4.15) est un schéma qui s'écrit sous la forme

$$(\mathscr{S}) \quad \left\{ \begin{array}{lcl} x_{n+1} & = & x_n + h\phi(t_n,x_n,h), \quad \text{pour} \;\; n = 0, ..., N-1, \\ x(t_0) & = & x_0, \end{array} \right. \tag{4.16}$$

où la fonction ϕ sera à définir suivant le schéma choisi.

Exemple *Par exemple, pour le schéma d'Euler explicite nous avons*

$$\phi(t_n,x_n,h) = f(t_n,x_n)$$

4.2.3 Consistance, stabilité et convergence

Nous pouvons alors introduire les notions fondamentales de ce chapitre. Commençons par la consistance.

Définition 13 (ERREUR DE CONSISTANCE)

On appelle erreur de consistance, que l'on note τ_n, le nombre réel défini par

$$\tau_{n+1}(h) = x(t_{n+1}) - x(t_n) - h.\phi(t_n, x(t_n), h), \text{ pour tout } n = 0, ..., N-1.$$

Remarque
Une autre façon d'écrire l'erreur de consistance est comme ceci :

$$x(t_{n+1}) = x(t_n) + h.\phi(t_n, x(t_n), h) + \tau_{n+1}(h), \textit{ pour tout } n = 0, ..., N-1.$$

Ca s'interprète assez simplement. L'erreur de consistance τ_{n+1} est l'erreur que l'on ferait si l'on partait de la solution $x(t_n)$ comme condition initiale et que l'on ne faisait qu'une seule itération. Nous voyons ainsi ce que chacun des pas séparément peut engendrer individuellement comme erreur en supposant que l'on n'a pas fait d'approximation dans les pas précédents.

Définition 14 (SCHÉMA CONSISTANT)

On dit que le schéma explicite à un pas $(\mathscr{S})$ donné par (4.16) est consistant avec le problème de Cuachy $(\mathscr{C})$ si et seulement si

$$\lim_{h\to 0} \sum_{n=1}^{N} |\tau_n(h)| = 0$$

Une fois l'erreur de consistance définie, nous nous intéressons à comment les erreurs peuvent s'accumuler. Il se trouve en effet que les ordinateurs sont obligés de faire des arrondis. Par exemple, si la solution trouvée est $1/3$, l'ordinateur va arrondir à $0,3333333$. Et même avec une précision assez fine, il aura quand même perturbé légèrement le calcul précédent. La question que l'on se pose est la suivante :
si le schéma est sensible à la moindre perturbation, les approximations risquent de ne pas être maîtrisée du tout. Il faut donc trouver un critère permettant de vérifier que ces petites perturbations sont sous contrôle. Comment ? Grâce à la notion de stabilité.

Définition 15 (SCHÉMA STABLE)

On dit que le schéma explicite à un pas ($\mathscr{S}$) donné par (4.16) est stable s'il existe un réel $M \geq 0$ tel que, étant donnés $\varepsilon_1, ..., \varepsilon_N$, si l'on considère pour le schéma explicite

$$x_{n+1} = x_n + h + \phi(t_n, x_n, h), \text{ avec } x_0 \text{ donné,}$$

une perturbation de ce schéma que l'on note

$$y_{n+1} = y_n + h + \phi(t_n, x_n, h) + \varepsilon_{n+1}, \text{ avec } y_0 = x_0,$$

alors

$$\max_{0 \leq n \leq N} \|y_n - x_n\| \leq M \sum_{n=1}^{N} |\varepsilon_n|.$$

Autrement dit, si les erreurs d'approximation à chaque pas de temps ne sont pas très grandes, l'erreur pour la solution approchée au pas suivant reste maîtrisée.
Nous pouvons enfin introduire la troisième et dernière notion fondamentale : la notion de convergence.

Définition 16 (CONVERGENCE)

On dit que le schéma explicite à un pas ($\mathscr{S}$) donné par (4.16) est convergent vers la solution du problème de Cauchy ($\mathscr{C}$) si

$$\lim_{h \to 0} \max_{1 \leq n \leq N} \|x(t_n) - x_n\| = 0.$$

Et le théorème permettant de relier les trois notions.

Théorème 5 (THÉORÈME DE LAX)

Si le schéma explicite à un pas ($\mathscr{S}$) donné par (4.16) est stable et consistant avec le problème de Cauchy ($\mathscr{C}$), alors il converge vers la solution du problème de Cauchy ($\mathscr{C}$).

Le problème reste de vérifier que le schéma est stable et consistant. Mais en utilisant la définition seulement, l'exercice peut très vite s'avérer difficile. Il faudrait trouver des critères plus faciles à manipuler pour obtenir ces deux notions sans trop de difficulté. C'est tout le but de ce qui suit. Commençons par un critère simple pour vérifier la consistance.

Proposition 5 (CRITÈRE DE CONSISTANCE)

Un schéma explicite à un pas $(\mathscr{S})$ donné par (4.16) est consistant avec le problème de Cauchy $(\mathscr{C})$ si et seulement si
$$\phi(t,x,0) = f(t,x).$$

Puis un critère simple pour la stabilité.

Proposition 6 (CRITÈRE DE STABILITÉ)

S'il existe un réel $L > 0$ tel que pour tout $t \in [t_0, t_0 + T]$, pour tous x, $y \in \mathbb{R}$ et pour tout $h \leq T$,
$$\|\phi(t,x,h) - \phi(t,y,h)\| \leq L\|x - y\|,$$
alors le schéma explicite à un pas $(\mathscr{S})$ donné par (4.16) est stable et $M = e^{LT}$.

Quand un système est convergent, il nous reste à savoir maintenant à quelle vitesse il converge. Pour cela on introduit la notion d'ordre. L'ordre de la convergence est liée à l'ordre de la consistance, et plus l'ordre est élevé, plus le schéma converge rapidement vers la solution exact. Voyons ça d'un peu plus près.

Définition 17 (ORDRE DE CONSISTANCE)

On dit que le schéma explicite à un pas $(\mathscr{S})$ donné par (4.16) est consistant d'ordre p avec le problème de Cauchy $(\mathscr{C})$, s'il existe un réel $K > 0$ tel que

$$\sum_{n=1}^{N} |\tau_n(h)| \leq Kh^p.$$

Proposition 7 (ORDRE DE CONVERGENCE)

Si un schéma explicite à un pas $(\mathscr{S})$ donné par (4.16) est stable et consistant d'ordre p avec le problème de Cauchy $(\mathscr{C})$ alors

$$\max_{1 \leq n \leq N} \|x(t_n) - x_n\| \leq MKh^p.$$

On dit alors qu'il est convergent d'ordre p.

Là encore utiliser la définition peut s'avérer assez compliqué. Il faudrait un critère permettant de montrer l'ordre de consistance sans trop de difficulté.

Proposition 8 (ORDRE DE CONVERGENCE :CRITÈRES)

Pour toute fonction f de classe $\mathscr{C}^p$, le schéma explicite à un pas $(\mathscr{S})$ donné par (4.16) est consistant d'ordre au moins p avec le problème de Cauchy $(\mathscr{C})$ si

$$\begin{array}{rcl} \phi(t,x,0) & = & f(t,x), \\ \frac{\partial}{\partial h}\phi(t,x,0) & = & \frac{1}{2}Df(t,x) = \frac{1}{2}[\frac{\partial}{\partial t}f(t,x) + f(t,x)\frac{\partial}{\partial x}f(t,x)], \\ & \vdots & \\ \frac{\partial^{p-1}}{\partial h^{p-1}}\phi(t,x,0) & = & \frac{1}{p}D^{p-1}f(t,x), \end{array}$$

avec $D^{p-1}f(t,x) = D[D^{p-2}f(t,x)]$.

Exemple *La différentielle seconde de f est donnée en (t,x) par :*

$$D^2f = D[Df] = \frac{\partial^2}{\partial t^2}f + 2f\frac{\partial^2}{\partial x \partial t}f + \frac{\partial}{\partial t}f\frac{\partial}{\partial x}f + f(\frac{\partial}{\partial x}f)^2 + f^2\frac{\partial^2}{\partial x^2}f.$$

Exemple *On peut montrer que si f satisfait les hypothèse de l'existence et l'unicité globale, alors le schéma d'Euler explicite correspondant au problème de Cauchy $(\mathscr{C})$ est convergent d'ordre exactement 1.*

Exemple

Le schéma du point milieu :

le schéma du point milieu consiste à utiliser la méthode des rectangles, non plus à gauche (ni à droite d'ailleurs), mais au point milieu de chaque intervalle de la forme $[t_n, t_{n+1}]$. Il est donné par l'équation suivante

$$x_{n+1} = x_n + hf(t_n + \frac{h}{2}, x_n + \frac{h}{2}f(t_n, x_n)), \textit{ pour } n = 0, ..., N-1, \textit{ avec } x_0 \textit{ donné.}$$

On montre que ce schéma est exactement d'ordre 2.

Le schéma de Heun (méthode des trapèzes) :

le schéma de Heun est donné par le système suivant

$$\left\{ \begin{array}{rcl} \hat{x}_n & = & x_n + hf(t_n, x_n), \\ x_{n+1} & = & x_n + \frac{h}{2}\left(f(t_n, x_n) + f(t_{n+1}, \hat{x}_n)\right), \end{array} \right.$$

où $n = 0, ..., N-1$ et x_0 est donné.

On voit dans la deuxième équation que l'on estime la moyenne aux extrémités du segment $[t_n, t_{n+1}]$, ce qui correspond bien à l'approximation de l'aire sous l'intégrale par un trapèze. On peut écrire ce système autrement de la façon suivante

$$x_{n+1} = x_n + \frac{h}{2}\left(f(t_n, x_n) + f(t_n + h, x_n + hf(t_n, x_n))\right),$$

où $n = 0, ..., N-1$ *et* x_0 *est donné.*
On montre que ce schéma est exactement d'ordre 2.

FIGURE 4.2 – Karl Heun (1859 – 1929), mathématicien allemand qui à qui l'on doit le schéma numérique qui porte son nom basé sur la méthode des trapèzes.

Remarque
Le schéma explicite à un pas est convergent d'ordre exactement p *quand il est d'ordre au moins* p *et qu'il n'est pas d'ordre au moins* $p+1$.

4.2.4 Les méthodes de Runge-Kutta

Certains schémas comme les schémas du point milieu et de Heun (méthode des trapèzes) permettent d'obtenir des convergences d'ordre plus élevé qu'Euler explicite.

1. Dans le cas du schéma du point milieu, la dérivée en l'instant intermédiaire entre t_n et t_{n+1} est estimée.

2. Dans le cas du schéma de Heun, la moyenne entre les estimations effectuées aux instant t_n et t_{n+1} est estimée.

La méthode du point milieu est due à Carl Runge, puis Martin Kutta a proposé diverses méthodes basées sur des moyennes à différents instants.
Cette idée a ensuite été généralisée pour construire ce que l'on appelle désormais les schémas de Runge-Kutta de n'importe quel ordre.

Définition 18 (SCHEMA DE RUNGE-KUTTA)

La structure générale d'un schéma de Runge-Kutta à s-stages explicite est donné par le système suivant

$$(RK)\quad \left\{ \begin{array}{rcl} X_i & = & x_n + h \displaystyle\sum_{j=1}^{i-1} a_{ij} f(t_n + c_j h, X_j), \quad i = 1, ..., s \\ x_{n+1} & = & x_n + h \displaystyle\sum_{i=1}^{s} b_i f(t_n + c_i h, X_i). \end{array} \right. \tag{4.17}$$

avec $n = 0, ..., N-1$ et x_0 donné. Par convention, nous notons $\sum_{i=1}^{0} ... = 0$

La donnée d'une méthode de Runge-Kutta explicite est donc la donnée de :

1. s instants compris entre t_n et t_{n+1} définis par $t_n + c_i h$, où $c_i \in [0, 1]$, pour $i = 1, ..., s$,
2. s poids b_i, $i = 1, ..., s$ permettant de faire la moyenne pondérée finale,
3. poids intermédiaires a_{ij} permettant de faire les estimations intermédiaires.

De façon usuelle, on représente alors une méthode de Runge-Kutta grâce au tableau de Butcher :

$$\begin{array}{c|cccccc} c_1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ c_2 & a_{21} & 0 & 0 & \ldots & 0 & 0 \\ c_3 & a_{31} & a_{32} & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ c_{s-1} & a_{s-1,1} & a_{s-1,2} & a_{s-1,3} & \ldots & 0 & 0 \\ c_s & a_{s,1} & a_{s,2} & a_{s,3} & \ldots & a_{s,s-1} & 0 \\ \hline & b_1 & b_2 & b_3 & \ldots & b_{s_1} & b_s \end{array}$$

FIGURE 4.3 – John Butcher (1933 –), mathématicien Néo-Zélandais, spécialisé dans les méthodes numériques de résolution d'équations différentielles. On lui doit le tableau qui porte son nom pour décrire les schémas de Runge-Kutta.

Exemple

1. *Le schéma représenté par*

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

est le schéma d'Euler explicite.

2. *Le schéma représenté par*

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ \hline & 0 & 1 \end{array}$$

   est le schéma du point milieu.

3. *Le schéma représenté par*

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array}$$

   est le schéma de Heun. Ce dernier schéma est également appelé parfois le schéma de Runge-Kutta 2, même si le schéma du point milieu est lui aussi un schéma d'ordre 2 de type Runge-Kutta.

4. *Quand on parle "du" schéma de Runge Kutta, ou encore RK-4 on désigne le schéma à 4 stages suivant*

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array}$$

La question que l'on peut se poser est pourquoi ce choix des poids et des instants intermédiaires ?
La réponse est assez simple : pour avoir les propriétés de stabilité et de consistance d'ordre élevé.
Les schémas de Runge-Kutta explicites rentrent en effet dans le cadre général des schémas explicites à 1 pas puisqu'on peut les écrire sous la forme

$$\left\{ \begin{array}{lcl} x_{n+1} & = & x_n + h\phi(t_n, x_n, h), \quad \text{pour} \;\; n = 0, ..., N-1, \\ x(t_0) & = & x_0. \end{array} \right.$$

Afin d'étudier leurs propriétés, on peut donc utiliser la théorie générale des schémas explicites à un pas. Nous avons alors le résultat suivant.

Proposition 9 (CRITÈRES DE CONSISTANCE DE RUNGE-KUTTA)

On considère un schéma de Runge-Kutta sous sa forme générale (RK) donné par (4.17).

1. Ce schéma est consistant d'ordre au moins 1 avec le problème de Cauchy $(\mathscr{C})$ si et seulement si
$$\sum_{i=1}^{s} b_i = 1.$$
2. Ce schéma est consistant d'ordre au moins 2 avec le problème de Cauchy $(\mathscr{C})$ si et seulement si
$$\sum_{i=1}^{s} b_i c_i = \frac{1}{2} \quad \text{et} \quad \sum_{i=1}^{s} b_i \sum_{j=1}^{i-1} a_{ij} = \frac{1}{2}.$$

Remarque

1. *La condition* $\sum_{i=1}^{s} b_i = 1$ *reflète bien le fait que les* b_i *sont bien des coefficients d'une moyenne pondérée.*
2. *On obtient des relations du même type pour tous les ordres. Et on peut ainsi montrer que le schéma RK-4 est consistant d'ordre 4.*

Nous avons également un critère de stabilité pour les schémas de Runge-Kutta.

Proposition 10 (CRITÈRES DE STABILITÉ DE RUNGE-KUTTA)

Tous les schémas du type Runge-Kutta explicites sont stables dès que f est continue et globalement lipschitzienne par rapport à sa variable x sur l'intervalle $[t_0, t_0 + T]$ étudié.

Remarque

En particulier le schéma RK-4 est convergent d'ordre 4. C'est le schéma le plus communément utilisé. Il est à la base du solveur ODE45 de Matlab. Le solveur ODE45 est à pas de temps adaptatif (à chaque instant, le pas de temps h_n *n'est pas prédéfini, on le détermine pour que l'erreur soit inférieure à une précision donnée).*
Pour ça, on fait une itération de RK-4 que l'on compare avec une itération d'un schéma de Runge-Kutta d'ordre 5 (dû à Dormand et Prince en 1980).
La différence $|RK - 4 - RK5|$ *donne alors une estimation de l'erreur par RK-4. On fait donc comme si RK-5 donnait la solution exacte. Si l'erreur est supérieure à la précision que l'on a choisie préalablement, on diminue le pas de temps* h_n.

Les méthodes de Runge-Kutta sont donc basées sur l'introduction d'instants intermédiaires entre t_n et t_{n+1} pour augmenter l'ordre de la consistance.
Un autre moyen d'augmenter la précision du calcul consiste à faire intervenir plusieurs pas de temps lors du calcul de x_{n+1}. C'est ce qu'on appelle les schémas à pas multiples ou multipas explicites. Ces schémas sont de la forme

$$x_{n+1} = x_n + h\phi(t_n, x_n, h_n, t_{n-1}, x_{n-1}, h_{n-1}, t_{n-2}, x_{n-2}, h_{n-2}, ...).$$

Ils sortent évidemment du cadre des schémas à un pas, mais on peut pour elles aussi définir la convergence que l'on décompose en stabilité et consistance. Le schéma explicite multipas le plus couramment utilisé est celui d'Adams-Bashforth (dû à John Adams qui l'utilisa en 1977 pour résoudre une EDO décrivant un problème de capillarité introduit par Francis Bashforth en 1883). Les méthodes multipas sont préférables aux méthodes de Runge-Kutta lorsque la fonction est coûteuse (en temps de calcul) à évaluer.
Mais toutes ces méthodes, que ce soit Runge-Kutta explicite ou multipas explicites, ont un point commun : elles ne sont performantes que pour les problèmes non-raides (en anglais Non STIFF).

4.3 Problèmes raides et schémas implicites

Lorsque l'on applique un schéma explicite à un pas adaptatif (comme ODE45) à certains problèmes dont la solution présente des problèmes raides. On obtient un effet en accordéon.
Le programme doit placer beaucoup de points dans les problèmes raides pour obtenir une précision requise.
Si le problème est trop raide, le nombre de points nécessaires devient trop important, le temps de calcul devient trop long et le schéma est inefficace.
La question est alors la suivante : comment caractérise-t-on les problèmes raides ? Ce sont les zones où la constante de Lipschitz L est trop grande.
Ceci vient du fait que l'erreur d'un schéma explicite stable et d'ordre au moins p est

$$\max_{1\leq n\leq N} \|x(t_n) - x_n\| \leq MKh^p,$$

où K est la constante de consistance, et $M = e^{LT}$ la constante de stabilité.
Ainsi, lorsque L devient grande, M devient grande exponentiellement, et donc la valeur prise par h doit être très petite pour garder une précision initialement donnée.
A cause de ça, on introduit une autre notion de stabilité qui teste le schéma sur les différentes valeurs possibles de L.

4.3.1 Test Linéaire Standard

Cette notion est basée sur ce que l'on appelle l'équation Test Linéaire Standard (ou TLS).
Ce test est défini par le système suivant

$$(TLS) \quad \left\{ \begin{array}{rcl} x'(t) & = & -Lx(t), \\ x(0) & = & x_0, \end{array} \right. \tag{4.18}$$

où $L > 0$.

1. La solution exacte est
$$x(t) = x_0 e^{-Lt}.$$
2. Elle vaut x_0 pour $t = 0$ et tend vers 0 quand t tend vers l'infini.
3. D'un autre côté, plus L est grand, plus $x(t)$ tend vers 0 de façon raide.
4. Enfin, nous remarquons que la constante de Lipschitz du problème (TLS) est L. En effet, si l'on pose $f(t, x) = -Lx$ pour $t \in \mathbb{R}_+$ et $x \in \mathbb{R}$, alors pour tous $t \in \mathbb{R}_+$ et pour tous $x, y \in \mathbb{R}$,
$$|f(t, x) - f(t, y)| = |-Lx + Ly| = L|x - y|.$$

Définition 19 (SCHEMA A-STABLE)

On dit qu'un schéma est A-stable si et seulement si ce schéma appliqué au problème (TLS) donne une solution x_n vérifiant $\lim_{n \to +\infty} x_n = 0$ quel que soit $L > 0$, et quel que soit le pas de temps constant h.
On dit aussi que le schéma est **inconditionnellement stable**.

Remarque

1. *Un schéma est A-stable est donc un schéma qui peut traiter les problèmes de n'importe quelle raideur sans condition sur le pas de temps.*
2. *La stabilité d'un schéma décrit la façon dont les erreurs s'accumulent sur un intervalle de temps borné $[t_0, t_0 + T]$ tandis que la A-stabilité décrit la façon dont les erreurs faussent le comportement de la solution pour $t \to +\infty$.*

Exemple *On peut montrer que le schéma d'Euler explicite n'est pas A-stable.*

De même on peut montrer que tous les schémas de Runge-Kutta explicites ont des conditions sur le pas de temps dépendant de la raideur du problème. Ces schémas ne sont donc pas A-stables.
Pour cette raison, on a introduit d'autres schémas qui cette fois-ci seront A-stables. En contre partie, ils sont plus difficiles à analyser et à coder. Le plus simple de ces schémas est le schéma d'Euler implicite.

4.3.2 Schéma d'Euler implicite (ou rétrograde)

Le schéma d'Euler implicite est construit de la même manière que le schéma d'Euler explicite : par la méthode des rectangles. Mais au lieu de considérer le rectangle à gauche, on considère le rectangle à droite cette fois-ci.
Le schéma s'écrit alors de la façon suivante :

$$(\mathscr{E}_r) \quad \left\{ \begin{array}{lcl} x_{n+1} & = & x_n + hf(t_n + h, x_{n+1}), \quad \text{pour } n = 0, ..., N-1, \\ x(t_0) & = & x_0. \end{array} \right.$$

Ce schéma est appelé implicite parce qu'il ya du x_{n+1} dans le second membre. A chaque itération, le schéma définit donc une équation qu'il faut résoudre pour trouver x_{n+1} de façon implicite.

Proposition 11 (SCHEMA EULER IMPLICITE A-STABLE)

Les schéma d'Euler implicite est A-stable.

4.3.3 Shémas de Runge-Kutta implicite

Définition 20 (SCHEMAS RK IMPLICITES)

Un schéma de Runge-Kutta implicite à S-stages est décrit par le tableau de Butcher suivant

$$\begin{array}{c|cccccc} c_1 & a_{11} & a_{12} & a_{13} & \dots & a_{1,s-1} & a_{1s} \\ c_2 & a_{21} & a_{22} & a_{23} & \dots & a_{2,s-1} & a_{2,s} \\ c_3 & a_{31} & a_{32} & a_{33} & \dots & a_{3,s-1} & a_{3,s} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ c_{s-1} & a_{s-1,1} & a_{s-1,2} & a_{s-1,3} & \dots & a_{s-1,s-1} & a_{s-1,s} \\ c_s & a_{s,1} & a_{s,2} & a_{s,3} & \dots & a_{s,s-1} & a_{s,s} \\ \hline & b_1 & b_2 & b_3 & \dots & b_{s_1} & b_s \end{array}$$

ou par la forme développé suivante

$$(RK_r) \quad \left\{ \begin{array}{rcl} X_i & = & x_n + h \displaystyle\sum_{j=1}^{s} a_{ij} f(t_n + c_j h, X_j), \quad i = 1, ..., s \\ x_{n+1} & = & x_n + h \displaystyle\sum_{i=1}^{s} b_i f(t_n + c_i h, X_i). \end{array} \right. \tag{4.19}$$

avec $n = 0, ..., N-1$ et x_0 donné.

On pourrait montrer que tous les schémas de implicite à S-stages sont A-stables.
Le cas explicite est d'ailleurs un cas particulier du cas implicite où le tableau de Butcher contient une matrice "strictement triangulaire inférieure".

Exemple
Le schéma représenté par

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

est le schéma d'Euler implicite.

La convergence pour les schémas de Runge-Kutta implicites se définit comme pour les schémas explicites.

$$\lim_{h\to 0} \max_{1\leq n\leq N} \|x(t_n) - x_n\| = 0.$$

La notion de consistance et d'ordre de consistance sont légèrement différentes par rapport au cas explicite, mais sont assez similaires quand même.
Les conditions sur l'ordre de consistance sont les mêmes à savoir :

1. Ce schéma est consistant d'ordre au moins 1 avec le problème de Cauchy ($\mathscr{C}$) si et seulement si
$$\sum_{i=1}^{s} b_i = 1.$$
2. Ce schéma est consistant d'ordre au moins 2 avec le problème de Cauchy ($\mathscr{C}$) si et seulement si
$$\sum_{i=1}^{s} b_i c_i = \frac{1}{2} \quad \text{et} \quad \sum_{i=1}^{s} b_i \sum_{j=1}^{s} a_{ij} = \frac{1}{2}.$$
mais la matrice constituée des a_{ij} n'est plus strictement triangulaire inférieure.

La convergence est toujours vraie pour les méthodes de Runge-Kutta implicites dès qu'elles sont consistantes et que f vérifie les conditions du théorème de Cauchy-Lipschitz global.

4.3.4 Résolution des itérations des schémas implicites

On a vu que les schémas implicites définissent à chaque itération une équation qu'il faut résoudre pour trouver x_{n+1}.

Exemple *Pour Euler implicite c'est*

$$x_{n+1} = x_n + hf(t_{n+1}, x_{n+1})$$

C'est à dire que x_{n+1} est un point fixe de la fonction G définie pour tout $x \in \mathbb{R}$ par

$$G(x) = x_n + hf(t_{n+1}, x).$$

il existe une théorie classique de l'existence et l'unicité des points fixes des fonctions (c'est le chapitre 3 de ce cours). Elle donne une condition suffisante d'existence et d'unicité de x_{n+1}.
Mais pour un problème raide, cette condition n'a que très peu d'intérêt car il faut par exemple pour le problème de Cauchy, si f est L-lipschitzienne, que $h < 1/L$ pour avoir l'existence et l'unicité du point fixe x_{n+1}.
Autrement dit, il faut une condition sur la valeur du pas de temps.
Pour des problèmes raides, cette condition est donc peu utile.
D'autres conditions sont actuellement développées qui donnent l'existence et l'unicité de x_{n+1} avec une condition moins restrictive sur h. Mais

1. il y a quand même une condition sur h,
2. ça ne marche pas pour tous les problèmes raides.

toutefois, en pratique, la solution x_{n+1} est en général bien définie pour les schémas de Runge-Kutta implicites, même pour des problèmes très raides et de grands pas de temps.
Enfin, quand x_{n+1} est bien définie, la méthode la mieux adaptée pour la calculer est la méthode de Newton que nous avons vus au chapitre précédent.

Bibliographie

[1] S. BENZONI-GAVAGE, *Calcul différentiel et équations différentielles*, Dunod, 2014. 41

[2] R.-L. BURDEN ET J.-D. FAIRES, *Numerical Analysis*, Brooks Cole, 2001. 3, 16, 26, 41

[3] J.-P. DEMAILLY, *Analyse numérique et équations différentielles*, PUG Grenoble, 1996. 41

[4] F. FILBET, *Analyse numérique : algorithme et étude mathématique*, Dunod, 2009. 3, 16, 26, 41

[5] R. QUARTERONI, A. ET SACCO ET F. SALERI, *Méthodes numériques pour le calcul scientifique. Programmes en MATLAB*, Springer, 2005. 3, 16, 26, 41

[6] M. SCHATZMAN, *Analyse numérique*, InterEditions, Paris, 1991. 3, 16, 26, 41

Index