MODULE MA322
CORRIGÉ DE TRAVAUX DIRIGÉS N° 4
———-
Résolution numérique des équations différentielles

Aéro. 3
Semestre : 2
A.U. : 2022-2023
Prof. H. El-Otmany

Exercice n°1 On considère un problème $(\mathcal{P})$ de Cauchy $y' = f(t, y)$ avec $y(a) = \alpha$ pour $t \in [a; a+T]$. La méthode de Heun est une méthode numérique à un pas où le calcul de y_{n+1} à partir de y_n est décrite par

$$\begin{cases} \text{(prédicteur)} & \bar{y}_n &= y_n + hf(t_n, y_n), \\ \text{(correcteur)} & y_n &= y_n + \frac{h}{2}\left(f(t_n, y_n) + f(t_{n+1}, \bar{y}_{n+1})\right). \end{cases}$$

1. Expliquer en quoi on peut affirmer que cette méthode est inspirée de la méthode des trapèzes.
D'après le problème $(\mathcal{P})$ de Cauchy, la solution exacte $y(t)$ vérifie $y' = f(t, y)$. Ce qui nous donne par intégration entre t_n et t_{n+1} :

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

Supposons que le pas de subdivision (ou discrétisation) $h_n = t_{n+1} - t_n$ de l'intervalle $I = [a; a+T]$ est constant et vaut $h = \frac{(a+T)-a}{N} = \frac{T}{N}$, $N \in \mathbb{N}^*$. (La généralisation à un pas non constant est triviale). En appliquant la méthode des Trapèzes pour approcher l'intégrale (1), on obtient

$$y(t_{n+1}) = y(t_n) + \frac{h}{2}\left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right) - \frac{h^3}{12} y^{(3)}(\xi), \quad \xi \in [t_n; t_{n+1}].$$

où le terme $y^{(3)}$ correspond à la dérivée seconde par rapport t de la fonction $f(t, y(t))$. Ceci mène au schéma de Crank-Nicolson :

$$y(t_{n+1}) = y(t_n) + \frac{h}{2}\left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right), \quad 0 \leqslant n \leqslant N-1.$$

qui est implicite comme le schéma d'Euler implicite (ou rétrograde).
On peut construire un nouveau schéma dit prédicteur-correcteur d'Euler-Cauchy (Heun) via les étapes suivantes :
— on cherche d'abord une estimation grossière de y_{n+1}, notée $\bar{y}_{n+1}$, via l'utilisation par exemple de la méthode d'Euler explicite
— on améliore ensuite cette estimation en s'inspirant du schéma d'Euler implicite.
— Enfin, on établit le schéma

$$\begin{cases} \text{(prédicteur)} & \bar{y}_n &= y_n + hf(t_n, y_n), \\ \text{(correcteur)} & y_{n+1} &= y_n + \frac{h}{2}\left(f(t_n, y_n) + f(t_{n+1}, \bar{y}_n)\right), \quad 0 \leqslant n \leqslant N-1. \end{cases}$$

Remarques : Les schémas à un pas explicites ou prédicteur-correcteur D'Euler-Cauchy peuvent s'écrire sous la forme générique

$$\begin{cases} y_0 &= y(t_0) = y(a) \\ y_{n+1} &= y_n + h\phi(t_n, y_n, h), \quad 0 \leqslant n \leqslant N-1, h = \frac{T}{N} \end{cases}$$

où ϕ est supposée de classe C^1 (cependant y est supposée de classe C^1).

— L'erreur de consistance (locale) à l'instant n est définie comme l'erreur commise par la solution exacte dans la méthode numérique :

$$\varepsilon_n = y(t_{n+1}) - y(t_n) - h\phi(t_n, y(t_n, h).$$

— L'erreur (locale) due à la méthode numérique de la méthode à un pas est définie par $e_{n+1} = y(t_{n+1}) - y_{n+1}$ où $y(t_{n+1})$ et y_{n+1} sont respectivement la solution exacte et le schéma numérique vérifiant

$$\begin{aligned} y(t_{n+1}) &= y(t_n) + h\phi(t_n, y(t_n), h) + \varepsilon_n, \\ y_{n+1} &= y_n + h\phi(t_n, y_n, h). \end{aligned}$$

— La méthode est dite consistante si pour toute solution du problème de Cauchy on a $\lim\limits_{h\longrightarrow 0} \sum\limits_{n=1}^{N} |\varepsilon_n| = 0$ ou bien $\lim\limits_{h\longrightarrow 0} \max\limits_{0\leqslant n\leqslant N-1} |\dfrac{y(t_{n+1}) - y_{n+1}}{h}| = 0$ (toute méthode d'ordre strictement positif est consistante).

— La méthode est dite d'ordre p s'il existe une constante $C > 0$ telle que $\sum\limits_{n=1}^{N} |e_n| \leqslant C\, h^p$ ou bien $\max\limits_{0\leqslant n\leqslant N-1} |\dfrac{y(t_{n+1}) - y_{n+1}}{h}| \leqslant C\, h^p$ (en pratique, on utilise $|\varepsilon_n| \leqslant Ch^{p+1}$).

— La méthode est dite convergente si $\lim\limits_{h\longrightarrow 0} \max\limits_{0\leqslant n\leqslant N-1} |y(t_{n+1}) - y_{n+1}| = 0$ (La consistance n'implique pas la convergence).

— La méthode est dite stable s'il existe une constante $L > 0$ telle que

$$\forall t \in [a, a+T],\ \forall y, z \in \mathbb{R},\ \forall h \in [0, h^*],\ |\phi(t, y, h) - \phi(t, z, h)| \leqslant L|y - z|.$$

— Toute méthode consistante et stable est convergente.

2. Montrer que cette méthode est consistante.

$$\phi(t, y, h) = \frac{1}{2}f(t, y) + \frac{1}{2}f(t + h, y + hf(t, y)),$$

On en déduit que $\phi(t, y, 0) = f(t, y)$. on utilise le développement de ε_n en puissance. de h. On a, avec $y'(t_n) = f(t_n, y(t_n))$,

$$\begin{aligned} y(t_{n+1}) &= y(t_n) + hy'(t_n) + \frac{h^2}{2}y''(\xi_n) \\ \phi(t_n, y_n, h) &= \phi(t_n, y_n, 0) + h\frac{\partial\phi}{\partial k}(t_n, y_n, \xi) \end{aligned}$$

Donc

$$\begin{aligned} \varepsilon_n &= y(t_{n+1}) - y(t_n) - h\phi(t_n, y_n, h) \\ &= h\left(f(t_n, y(t_n)) - \phi(t_n, y(t_n), 0)\right) + \frac{h^2}{2}\left(y''(\xi_n) - 2\frac{\partial\phi}{\partial k}(t_n, y_n, \xi)\right) \end{aligned}$$

Pour que le schéma à un pas soit consistante, il faut que $\phi(t, y, 0) = f(t, y)$ pour tout $(t, y) \in [a; a+T] \times \mathbb{R}$. On obtient ainsi

$$\varepsilon_n \leqslant \frac{h^2}{2}\left|y''(\xi_n) - 2\frac{\partial\phi}{\partial k}(t_n, y_n, \xi)\right| \leqslant C\, h^2$$

Ici, nous avons utilisé le fait que f est de classe C^1 et ϕ de classe C^1.

Remarque : la méthode à un pas est dite consistante si et seulement si pour tout $(t, y) \in [a; a+T] \times \mathbb{R}$, on a $\phi(t, y, 0) = f(t, y)$. Dans l'exercice du partiel, vous pouvez directement appliquer cette remarque. Elle valable pour toute méthode numérique avec sa fonction associée ϕ.

3. Montrer que cette méthode est stable.
   D'après le schéma d'Heun, on a

$$\phi(t,y,h) = \frac{1}{2}f(t,y) + \frac{1}{2}f(t+h, y+hf(t,y)),$$

   Pour $y, z \in \mathbb{R}$, on a

$$\begin{aligned}\phi(t,y,h) - \phi(t,z,h) &= \frac{1}{2}f(t,y) + \frac{1}{2}f(t+h,y+hf(t,y)) - \frac{1}{2}f(t,z) - \frac{1}{2}f(t+h,z+hf(t,z)) \\ &= \frac{1}{2}\left[f(t,y) - f(t,z)\right] + \frac{1}{2}\left[f(t+h,y+hf(t,y)) - f(t+h,z+hf(t,z))\right]\end{aligned}$$

   Supposons que f vérifie la condition de Lipschitz ($|f(t,u)-f(t,v)| \leqslant k|u-v|$). Par application de la condition de Lipschitzienneté avec $u = x$ et $v = y$, puis avec $u = y+hf(t,)$ et $v = z+hf(t,z)$)

$$\begin{aligned}|\phi(t,y,h) - \phi(t,z,h)| &\leqslant \frac{k}{2}|y-z| + \frac{k}{2}|y+hf(t,y) - z - hf(t,z)| \\ &\leqslant \frac{k}{2}|y-z| + \frac{k}{2}|y-z| + \frac{k}{2}|hf(t,y) - hf(t,z)| \\ &\leqslant k|y-z| + \frac{hk^2}{2}|y-z| = \left(k + \frac{hk^2}{2}\right)|y-z|.\end{aligned}$$

   D'où, la fonction ϕ vérifie la condition de Lipschitz avec $L = k + \frac{hk^2}{2}$. Par conséquent, le schéma d'Heun est stable.
   $\leadsto$ La méthode d'Heun est stable et consistante, donc elle est convergence.

Exercice n°2 On étudie l'équation différentielle $y' = -y$. Cette équation sera considérée sur $[0;10]$ avec la condition initiale $y(0) = 1$.

1. Déterminer la solution exacte de ce problème de Cauchy. Vérifier qu'on a toujours

$$\forall t \in [0;10],\ 0 < y(t) \leqslant 1.$$

   On a $f(t,y) = -y$ est 1-Lipschitzienne et continue sur $[0;10]$. Ainsi, le théorème de Cauchy-Lipschitz nous assure que ce problème admet une unique solution y définie sur $[0;10]$.
   On a $y' = -y$ donc $\frac{y'}{y} = -1$ (si $y \neq 0$ pour tout $t \in [0;10]$), soit $\ln|y| = -t + k \in \mathbb{R}$. Par conséquent $y(t) = Ce^{-t}$ où C est une constante. Comme $y(0) = 1$ alors $C = 1$. On en déduit que $y(t) = e^{-t}$ est la solution du problème de Cauchy.
   On a $0 \leqslant t \leqslant 10$ donc $-10 \leqslant -t \leqslant 0$. En utilisant la monotonie de $t \longmapsto e^{-t}$, on obtient $0 < e^{-10} \leqslant e^{-t} \leqslant e^0 = 1$.
2. Exprimer pour la méthode d'Euler explicite, l'expression de y_n en fonction de n.
   Soit $y(t_n) \approx y_n$ tel que $t_n = nh$. Le schéma d'Euler explicite s'écrit ainsi $y_n = y_{n-1} + hf(t_{n-1}, y_{n-1})$ avec $y_0 = y(t_0)$. Comme $f(t,y) = -y$, le schéma d'Euler explicite s'écrit ainsi $y_n = (1-h)y_{n-1}$. Par récurrence, on a

$$y_n = (1-h)y_{n-1} = (1-h)^2 y_{n-2} = \cdots = (1-h)^n y_0.$$

3. Déterminer à quelle condition sur h, on peut assurer que pour tout n, $0 < y_n \leqslant 1$.
   Pour que, pour tout n : $0 < y_n \leqslant 1$, il faut que h vérifie y_n converge vers la solution exacte lorsque n est très grand. Ceci est réalisable pour les valeurs de $h < 1$.

4. Majorer sommairement $\max\limits_{0\leqslant n\leqslant N} |y_n - y(t_n)|$. (On se limitera au cas où $h \leqslant \frac{1}{2}$.) On note $e_n = y_n - y(t_n)$ l'erreur locale du schéma numérique et $\mathcal{E}(h) = \max\limits_{0\leqslant n\leqslant N} |e_n|$ l'erreur globale du schéma numérique. On a par définition de l'erreur de consistance

$$\varepsilon_n = y(t_{n+1}) - y_{n+1} = y(t_{n+1}) - y(t_n) - hf(t_n, y(t_n))$$

En utilisant $y(t_{n+1}) = y(t_n) + hy'(t_n) + \frac{h^2}{2}y''(\xi_n)$, $\xi_n \in [t_n, t_{n+1}]$ et $y'(t_n) = f(t_n, y(t_n))$, on arrive à

$$\varepsilon_n = \frac{h^2}{2}y''(\xi_n).$$

D'après la question 2, on en déduit que

$$|\varepsilon_n| \leqslant \frac{h}{2}|y''(\xi_n)| \leqslant \frac{h}{2}\sup_{t\in[0;T]}|y''(t)| \leqslant \frac{h^2}{2}.$$

Comme $e_n = y(t_n) - y_n$ et f est 1-Lipschitzienne en y uniformément en t, on obtient

$$e_{n+1} = e_n + hf(t_n, y(t_n)) - hf(t_n, y_n) + \varepsilon_n \leqslant (1+h)e_n + \varepsilon_n$$

D'où, pour tout $1 \leqslant n \leqslant N$, on a

$$|e_{n+1}| \leqslant (1+h)|e_n| + |\varepsilon_n| \leqslant (1+h)|e_n| + \frac{h^2}{2}$$

Lemme de Gronwall : Soit $(x_n)_{n\geqslant 0}$ une suite réelle positives vérifiant

$$x_{n+1} \leqslant \alpha x_n + \beta, \alpha > 0, \beta > 0.$$

Alors, pour tout $n \geqslant 0$, on a

$$x_n \leqslant \alpha^n x_0 + \beta\sum_{k=0}^{n-1}\alpha^k = \alpha^n x_0 + \beta\frac{1-\alpha^n}{1-\alpha}.$$

De plus, si $\alpha = 1+\gamma$ avec $\gamma > 0$, comme $(1+\gamma)^n \leqslant e^{n\gamma}$, on a pour tout $n \geqslant 0$

$$x_n \leqslant e^{n\gamma}x_0 + \frac{\beta}{\gamma}\left(e^{n\gamma} - 1\right).$$

En appliquant le lemme de Gronwall avec $\gamma = h$ et $\beta = \frac{h^2}{2}$, on a pour tout $1 \leqslant n \leqslant N-1$:

$$|e_n| \leqslant e^{nh}|e_0| + \frac{h}{2}\left(e^{nh} - 1\right)$$

Comme $nh \leqslant Nh = T$ pour $0 \leqslant n \leqslant N$, il vient que

$$\max_{0\leqslant n\leqslant N}|e_n| \leqslant e^T|e_0| + \frac{h}{2}\left(e^T - 1\right).$$

Ainsi, si $e_0 = 0$ alors

$$\max_{0 \leqslant n \leqslant N} |e_n| \leqslant \frac{h}{2}\left(e^T - 1\right).$$

Remarque : Pour majorer $\max\limits_{0 \leqslant n \leqslant N} |y_n - y(t_n)|$, vous pouvez utiliser directement les expressions explicites de y_n et $y(t_n)$ avec $t_n = nh \leqslant 10$. On a

$$|y_n - y(t_n)| \leqslant |(1-h)^n - e^{-nh}| \leqslant n|1 - h - e^{-h}| \leqslant \frac{nh}{2}h \leqslant 5h$$

Cette majoration peut ne pas donner des informations sur l'évolution d'erreur dans le domaine physique et vous ne pourriez pas détecter des artefacts ou des dispersions dans le domaine physique. je vous conseille d'éviter ce calcul pour déterminer les erreurs numériques pour un projet industriel.

5. Reprendre les questions 2 à 4 avec les autres méthodes vues (Euler implicite, Runge-Kutta d'ordres 2 et 4). Retrouve-t-on des résultats compatibles avec les ordres connus pour ces méthodes ?

— Schéma d'Euler implicite :

(a) Soit $y(t_n) \approx y_n$ tel que $t_n = nh$. Le schéma d'Euler explicite s'écrit ainsi $y_n = y_{n-1} + hf(t_n, y_n) = y_n - hy_n$ avec $y_0 = y(t_0)$. Comme $f(t,y) = -y$, le schéma d'Euler explicite s'écrit ainsi $y_n = \frac{1}{1+h}y_{n-1} = \cdots = \frac{1}{(1+h)^n}y_0 = \frac{1}{(1+h)^n}$.

(b) Il n'y a pas de conditions sur la pas de discrétisation pour les schémas implicites (ou rétrogrades).

(c) Il suffit d'utiliser la même démarche pour prouver la majoration de l'erreur d'ordre h.

— RK2 :

(a) $y_n = \left(1 - h + \frac{h^2}{2}\right)^n$.

(b) Il faut que $h < 2$ pour que $0 < y_n \leqslant 1$.

(c) Il suffit d'utiliser la même démarche pour prouver la majoration de l'erreur d'ordre h^2.

— RK4 :

(a) $y_n = \left(1 - h + \frac{h^2}{2} - \frac{h^3}{6} + \frac{h^4}{24}\right)^n$.

(b) Il faut que $h < 1$ pour que $0 < y_n \leqslant 1$.

(c) Il suffit d'utiliser la même démarche pour prouver la majoration de l'erreur d'ordre h^4.

Exercice n°3 On considère un problème de Cauchy $(\mathcal{P})$ suivant (avec $T > 0$) :

$$(\mathcal{P}) \begin{cases} \forall t \in [0;T],\ y' & = \sin(y(t)) + \sin(t), \\ y(0) & = 0. \end{cases}$$

1. Montrer que ce problème de Cauchy a une et une seule solution.
Posons $f(t,y) = \sin(y(t)) + \sin(t)$, le problème de Cauchy $(\mathcal{P})$ s'écrit ainsi

$$(\mathcal{P}) \begin{cases} \forall t \in [0;T],\ y' & = f(t, y(t)), \\ y(0) & = 0. \end{cases}$$

f est une fonction de classe C^1 (somme de fonctions de classe C^1) et on $\frac{\partial f(t,y)}{\partial y} = \cos(y)$, donc $\left|\frac{\partial f(t,y)}{\partial y}\right| = |cos(y)| \leqslant 1$. Par conséquent, f est une fonction $1-$lipschitzienne. Ainsi le théorème de Cauchy-Lipschitz nous assure que le problème différentiel $(\mathcal{P})$ admet une unique solution $y : I \longmapsto \mathbb{R}$.
On peut aussi démontrer que f est 1-Lipschitzienne en utilisant pour tout $x \in \mathbb{R}$, $\sin x \leqslant x$ (facile à vérifier pour $x \geqslant 1, \sin x \leqslant x$ et pour $0 \leqslant x < 1$ on traite la fonction $f(x) = \sin x - x$).

2. Majorer sommairement $|y''|$ sur $[0; T]$.
La fonction $y'(t) = f(t, y(t))$ est de classe C^1 sur $[0; T]$ (somme et composé de fonctions dérivables et continues) où $f(t, y(t)) = \sin(y(t)) + \sin(t)$. Par conséquent

$$\begin{aligned} y''(t) &= \frac{\partial f(t,y)}{\partial t} + \frac{\partial f(t,y)}{\partial y} y'(t) \\ &= y'(t)\cos(y(t)) + \cos(t) + \cos(y(t))y'(t) = 2y'(t)\cos(y(t)) + \cos(t) \end{aligned}$$

Par conséquent

$$|y''(t)| \leqslant |2y'(t)\cos(y(t)) + \cos(t) + \cos(t)| \leqslant 5$$

3. Déterminer une majoration de l'erreur $\max_{0 \leqslant n \leqslant N} |y(t_n) - y_n|$ pour la méthode d'Euler explicite.
On a par définition :

$$\varepsilon_n = y(t_{n+1}) - y_{n+1} = y(t_{n+1}) - y(t_n) - hf(t_n, y(t_n))$$

En utilisant $y(t_{n+1}) = y(t_n) + hy'(t_n) + \frac{h^2}{2} y''(\xi_n),\ \xi_n \in [t_n, t_{n+1}]$ et $y'(t_n) = f(t_n, y(t_n))$, on arrive à

$$\varepsilon_n = \frac{h^2}{2} y''(\xi_n).$$

D'après la question 2, on en déduit que

$$|\varepsilon_n| \leqslant \frac{h}{2}|y''(\xi_n)| \leqslant \frac{h}{2} \sup_{t \in [0;T]} |y''(t)| \leqslant \frac{5}{2}h^2.$$

Comme $e_n = y(t_n) - y_n$ et f est 1-Lipschitzienne en y uniformément en t, on obtient

$$e_{n+1} = e_n + hf(t_n, y(t_n)) - hf(t_n y_n) + \varepsilon_n \leqslant (1+h)e_n + \varepsilon_n$$

D'où, pour tout $0 \leqslant n \leqslant N-1$, on a

$$|e_{n+1}| \leqslant (1+h)|e_n| + |\varepsilon_n| \leqslant (1+h)|e_n| + \frac{5}{2}h^2$$

En appliquant le lemme de Gronwall avec $\gamma = h$ et $\beta = \frac{5}{2}h^2$, on a pour tout $1 \leqslant n \leqslant N-1$:

$$|e_n| \leqslant e^{nh}|e_0| + \frac{5h}{2}\left(e^{nh} - 1\right)$$

Comme $nh \leqslant Nh = T$ pour $0 \leqslant n \leqslant N-1$, il vient que

$$\max_{0\leqslant n\leqslant N} |e_n| \leqslant e^T|e_0| + \frac{5h}{2}\left(e^T - 1\right).$$

Ainsi, si $e_0 = 0$ alors

$$\max_{0\leqslant n\leqslant N} |e_n| \leqslant \frac{5h}{2}\left(e^T - 1\right).$$

Remarque : on a $\lim\limits_{h\longrightarrow 0} \max\limits_{0\leqslant n\leqslant N} |e_n| = 0$. Donc la méthode est consistante.

4. Pour $T = 1$, quel valeur de N choisir pour garantir un résultat correct avec une précision de 10^{-3} ? Même question pour $T = 10$.
   Pour répondre à cette question, on cherche N tel que

$$\max_{0\leqslant n\leqslant N} |e_n| \leqslant \frac{5T}{2N}\left(e^T - 1\right) \leqslant \varepsilon.$$

— $\varepsilon = 10^{-3}$ et $T = 1$ et $h = \frac{T}{N} = \frac{1}{N}$. On cherche ainsi N tel que

$$\frac{5}{2N}\left(e^1 - 1\right) \leqslant \varepsilon.$$

Soit $N \geqslant \frac{5}{2}\left(e^1 - 1\right) \times 10^3$. Il suffit donc de choisir $N = \lfloor \frac{5}{2}\left(e^1 - 1\right) \times 10^3 \rfloor$, soit ainsi

$$N = \lfloor 4295.70 \rfloor = 4296.$$

— $\varepsilon = 10^{-3}$ et $T = 10$ et $h = \frac{T}{N} = \frac{10}{N}$. On cherche ainsi N tel que

$$\frac{50}{2N}\left(e^{10} - 1\right) \leqslant \varepsilon.$$

Soit $N \geqslant \frac{50}{2}\left(e^{10} - 1\right) \times 10^3$. Il suffit donc de choisir $N = \lfloor \frac{50}{2}\left(e^{10} - 1\right) \times 10^3 \rfloor$, soit ainsi

$$N = \lfloor 550636644.87 \rfloor = 550636645.$$

Dans ce cas, N est très grand, on peut espérer qu'en pratique un nombre de subdivision acceptable avant pour converger vers la solution exacte du problème.