FSAB1104 : solution de l'examen de janvier 2013

1 Hervé s'emberlificote avec Heun

En faisant appel à la méthode de Heun[1], Hervé souhaite résoudre numériquement le problème :

$$\left\{\begin{array}{rcl} x'(t) &=& -10\ x(t) - 9\ y(t),\\ y'(t) &=& -9\ x(t) - 10\ y(t),\\ \\ x(0) &=& 2,\\ y(0) &=& 0. \end{array}\right.$$

1. Est-ce que le problème d'Hervé est stable ?
   Justifier votre réponse.

   On écrit le système sous forme matricielle : $\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \underbrace{\begin{bmatrix} -10 & -9 \\ -9 & -10 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$

   Le problème sera stable si les valeurs propres de la matrice jacobienne $\mathbf{A}$ *sont négatives. Il suffit donc simplement de calculer :*

$$\begin{array}{rcl} \det \begin{bmatrix} -10-\lambda & -9 \\ -9 & -10-\lambda \end{bmatrix} &=& 0 \\ & \downarrow & \\ (10+\lambda)^2 - 81 &=& 0 \\ \lambda^2 + 20\lambda + 19 &=& 0 \\ & \downarrow & \textit{En observant que } 324 = 81 \times 4 \textit{ :-)} \\ \lambda &=& \dfrac{-20 \pm \sqrt{400-76}}{2} \ = \ -10 \pm \sqrt{81} \ = \ -10 \pm 9 \end{array}$$

   On conclut finalement que <u>le problème est stable</u>, puisque $\lambda' = -1$ *et* $\lambda'' = -19$ *sont négatives.*

   Un très grand nombre d'étudiants se sont lamentablement plantés dans cette question pourtant élémentaire. Dans le rayon du prix des horreurs, citons juste les étudiants justifiant la stabilité car la matrice est négative... Ah bon, c'est quoi le signe d'une matrice ? Ils vous ont appris quoi en algèbre, Kouider, Michel, Roland et tous les autres...

[1] Pour rappel, la méthode de Heun est définie comme suit :

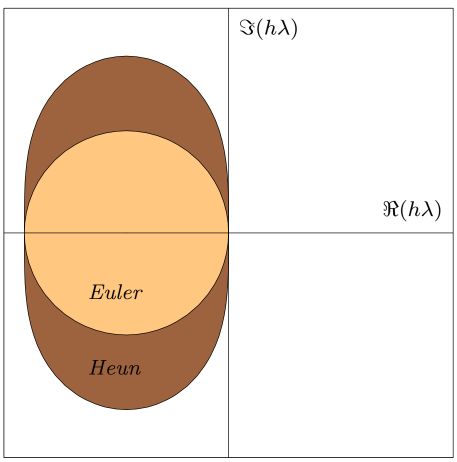
$$\begin{array}{rcl} U_{i+1} &=& U_i + \frac{h}{2}(K_1 + K_2) \\ \\ K_1 &=& f(X_i, U_i) \\ K_2 &=& f(X_i + h, U_i + hK_1) \end{array}$$

2. Donner l'expression analytique de la région de stabilité de la méthode de Heun.
   Esquisser cette région dans le plan complexe en définissant précisément les axes de la figure.
   Y indiquer aussi la zone de stabilité de la méthode d'Euler explicite.
   Ici, on considère évidemment le problème modèle habituel : $u'(t) = \lambda\ u(t)$.

   On évalue simplement U_{i+1} pour le problème modèle :

$$\begin{array}{rcl} K_1 &=& \lambda U_i \\ K_2 &=& \lambda\big(U_i + hK_1\big) \ = \ \lambda\big(U_i + h\lambda U_i\big) \\ & \Big\downarrow & \\ U_{i+1} &=& U_i + \dfrac{h}{2}\Big(K_1 + K_2\Big) = \left(1 + h\lambda + \dfrac{h^2\lambda^2}{2}\right) U_i \end{array}$$

La zone de stabilité de la méthode de Heun est donc définie par : $\boxed{\left|1 + h\lambda + \dfrac{h^2\lambda^2}{2}\right| \ \leq \ 1}$



Il est essentiel correctement d'indiquer les axes !

Non, la zone de stabilité de Heun n'a pas de petites oreilles comme la méthode de Runge-Kutta d'ordre 4. Noter qu'une très grande tolérance a été admise pour les dessins d'ellipse un brin approximative. Le plupart des horribles patates mal foutues a été admise :-)

Avoir les dessins des zones de stabilité sur son formulaire était parfaitement licite (et ici bien utile !)

3. Obtenir la contrainte de stabilité à imposer sur le pas de temps de Heun pour le problème d'Hervé.

   Il faut considérer la valeur propre la plus négative et écrire :

$$\begin{array}{rcccl} & & \left|1 - 19h + \dfrac{81h^2}{2}\right| & \leq & 1 \\ -1 & \leq & 1 - \dfrac{81}{2}h\left(\dfrac{2}{19} - h\right) & \leq & 1 \\ 2 & \geq & \dfrac{81}{2}h\left(\dfrac{2}{19} - h\right) & \geq & 0 \\ & & & \Big\downarrow & \\ & & h & \leq & \dfrac{2}{19} \end{array}$$

C'est exactement la même condition que celle obtenue avec la méthode d'Euler explicite. Il est donc possible d'obtenir immédiatement le résultat en observant le graphique, en n'effectuant aucun calcul, mais en se rappelant juste que pour une valeur propre réelle, la condition de stabilité pour Euler explicite était $h\lambda < 2$. C'est évidemment la valeur propre la plus raide qu'il faut considérer :-)

4. Ecrire une fonction MATLAB

```
function [x y] =  HeunHerve(n,dt)
```

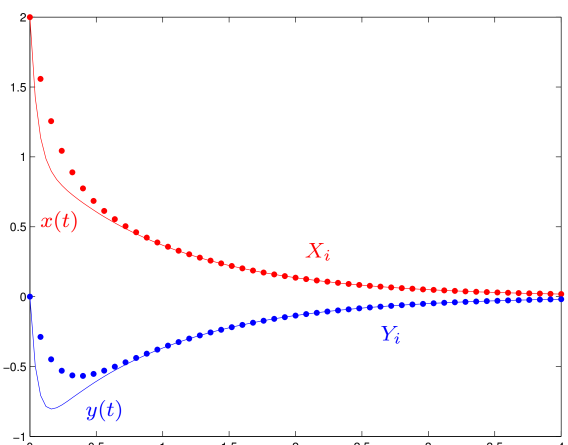
qui retourne deux vecteurs de taille `n+1` qui contiennent les différentes valeurs de $x(t)$ et $y(t)$ obtenues en effectuant n pas avec la méthode de Heun. On commence le calcul à partir de la condition initiale en $t = 0$. Le pas de temps est donné par le second argument `dt`.

Une implémentation possible est donnée par :

```
function [x y] =  HeunHerve(n,h)
X  = [[2 0] ; zeros(n,2)];
for k=1:n
    K1 = f(X(k,:));
    K2 = f(X(k,:) + h * K1);
    X(k+1,:) = X(k,:) + h * (K1 + K2)/2;
end
x = X(:,1); y = X(:2);
end

function dxdt = f(x)
dxdt = x;
dxdt(1) = -10*x(1) -9*x(2);
dxdt(2) =  -9*x(1)-10*x(2);
end
```

Comme la méthode de Heun pour un problème linéaire (ce qui est le cas ici !) est strictement équivalent à la méthode de Taylor, on peut aussi se contenter d'implémenter celle-ci (c'est encore plus simple et c'est évidemment admis :-). Par contre, effectuer d'abord tout le calcul pour `x` *et ensuite pour* `y` *est incorrect : il faut calculer les deux composantes du vecteur* `K1` *avant d'entamer le calcul de* `K2`*. Alors que ce programme est quasiment la copie conforme de celui d'un exercice effectué en séance tutorée, un nombre (vraiment !) très élevé d'étudiants n'arrivent pas à l'écrire...*



La solution analytique du problème d'Hervé est donnée par

$$\begin{cases} x(t) &= \exp(-19t) + \exp(-t) \\ y(t) &= \exp(-19t) - \exp(-t) \end{cases}$$

Même si cela ne faisait pas partie de l'examen, c'était vraiment facile de l'obtenir.

2 Stable ou pas stable ?

Pour résoudre numériquement une équation de transport

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x},$$

où c est une constante strictement positive, Francis considère le schéma numérique suivant :

$$\boxed{\frac{U_j^{n+1} - U_j^n}{\Delta t} \quad = \quad -c\ \frac{U_j^n - U_{j-1}^n}{\Delta x}}$$

où j et n sont respectivement les indices spatiaux et temporels.

On va effectuer l'analyse de stabilité en considérant l'évolution d'une pertubation quelconque de la forme $U_j^n = U^n \exp(ikX_j)$ et en calculant le carré du module $|G|^2$ du facteur d'amplification :

$$G = \frac{U^{n+1}}{U^n}.$$

1. Que représente k dans l'expression $U_j^n = U^n \exp(ikX_j)$?

   Le nombre entier k est le nombre d'onde (quelconque) de la perturbation (quelconque).
   Alors que l'enseignant pensait (naïvement) que cette question était vraiment facile, la plupart des étudiants n'ont pas été capable d'y répondre. Citons en vrac quelques perles particulièrement significatives : l'amplitude de la perturbation, le vecteur d'onde (c'est un scalaire : hein !), la direction de la propagation de l'onde, le coefficient de dilatation ou la conductivité thermique (sic !).

2. Ce schéma temporel est-il explicite ou implicite ?

   La méthode est explicite

3. Donner l'expression du nombre complexe G pour ce schéma.

   En incluant l'expression $U_j^n = U^n \exp(ikX_j)$ dans le schéma numérique, on obtient :

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = -c\ \frac{U_j^n - U_{j-1}^n}{\Delta x}$$

$$\Big\downarrow \ \textit{En remplaçant } U_j^n \textit{ par } U^n \exp(ikX_j)$$

$$\begin{aligned}
U^{n+1}\exp(ikX_j) - U^n\exp(ikX_j) &= -\frac{c\Delta t}{\Delta x}\Big(U^n\exp(ikX_j) - U^n\exp(ikX_j)\exp(-ik\Delta x)\Big) \\
U^{n+1} &= U^n\left(1 - \frac{c\Delta t}{\Delta x}\Big(1 - \exp(-ik\Delta x)\Big)\right)
\end{aligned}$$

On en déduit donc que : $$\boxed{G \quad = \quad \left(1 - \frac{c\Delta t}{\Delta x}\Big(1 - \exp(-ik\Delta x)\Big)\right)}$$

4. En déduire la condition de stabilité en termes de c, Δx et Δt pour ce schéma[2].

Pour obtenir la condition de stabilité, il faut effectuer le calcul carré du module de G et vérifier que ce module soit inférieur à l'unité.

$$|G|^2 = \left(1 - \frac{c\Delta t}{\Delta x}\Big(1 - \cos(-k\Delta x)\Big)\right)^2 + \left(\frac{c\Delta t}{\Delta x}\sin(-k\Delta x)\right)^2$$

En définissant $\beta = \dfrac{c\Delta t}{\Delta x}$ et en faisant judicieusement appel à quelques formules trigonométriques :-)

$$\begin{aligned}
&= \left(1 - 2\beta\sin^2\left(\frac{k\Delta x}{2}\right)\right)^2 + \left(2\beta\sin\left(\frac{k\Delta x}{2}\right)\cos\left(\frac{k\Delta x}{2}\right)\right)^2 \\
&= 1 - 4\beta\sin^2\left(\frac{k\Delta x}{2}\right) + 4\beta^2\sin^4\left(\frac{k\Delta x}{2}\right) + 4\beta^2\sin^2\left(\frac{k\Delta x}{2}\right)\cos^2\left(\frac{k\Delta x}{2}\right) \\
&= 1 - 4\beta\sin^2\left(\frac{k\Delta x}{2}\right) + 4\beta^2\sin^2\left(\frac{k\Delta x}{2}\right)\underbrace{\left(\sin^2\left(\frac{k\Delta x}{2}\right) + \cos^2\left(\frac{k\Delta x}{2}\right)\right)}_{=1} \\
&= 1 + 4\beta(\beta - 1)\sin^2\left(\frac{k\Delta x}{2}\right)
\end{aligned}$$

Maintenant, il est possible de déduire une condition de stabilité sur le pas de temps...

$$\begin{aligned}
|G| &\leq 1 \\
4\beta(\beta - 1)\sin^2\left(\frac{k\Delta x}{2}\right) &\leq 0 \\
\beta - 1 &\leq 0 \\
\beta &\leq 1
\end{aligned}$$

Le schéma est conditionnellement stable et la condition de stabilité est : $\boxed{\dfrac{c\Delta t}{\Delta x} \leq 1}$

Certains étudiants trouvent le cas critique en analysant quelques angles particulier : cette démarche est admise:-) De manière assez surprenante, cette question assez calculatoire (et que l'enseignant estimait difficile) a été relativement bien réussie... alors qu'expliquer ce qu'était le nombre d'onde (ce que l'enseignant estimait élémentaire) a posé problème à beaucoup d'étudiants.

[2]Il peut être utile d'utiliser les relations suivantes :

$$\begin{aligned}
1 - \cos(\alpha) &= 2\sin^2(\alpha/2) \\
e^{i\alpha} &= \cos(\alpha) + i\sin(\alpha)
\end{aligned}$$

3 Différences finies compactes

Christian a découvert le concept de différences finies compactes :

$$\boxed{\alpha U'_{i-1} + U'_i + \alpha U'_{i+1} \quad = \quad \beta \frac{U_{i+1} - U_{i-1}}{2h} + \gamma \frac{U_{i+2} - U_{i-2}}{4h} + \mathcal{O}(h^p)}$$

Cela lui permet d'obtenir une estimation très précise des valeurs $U'_i \approx u'(X_i)$ de la dérivée d'une fonction périodique à partir de ordonnées $U_i = u(X_i)$ aux abscisses $X_i = ih$ avec $i = 1 \ldots n$.

La périodicité implique évidemment $U_1 = U_n$ et $U'_1 = U'_n$.

Malencontreusement, un informaticien facétieux lui a chapardé les valeurs des 3 coefficients α, β et γ.

1. Quel est l'ordre de précision le plus élevé[3] que l'on aura avec les α, β et γ les plus adéquats ? Justifier brièvement.

   Comme la formule est parfaitement symétrique, il n'y aura que des termes d'erreurs pairs. En choisissant de manière astucieuse α, β et γ, on éliminera les termes en $\mathcal{O}(h^0)$, $\mathcal{O}(h^2)$ et $\mathcal{O}(h^4)$.

   Donc, on peut espérer -sauf miracle numérique- que : $\boxed{p \quad = \quad 6}$

2. Quelles relations doivent satisfaire α, β et γ pour obtenir la méthode la plus précise ?

   En ne gardant que les termes utiles, on écrit les développements en série de Taylor :

$$\begin{aligned}
U_{i\pm1} &= U_i \ \pm\ hU'_i \ +\ \ldots\ \pm\ \frac{h^3}{6}U_i^{(3)} \ +\ \ldots\ \pm\ \frac{h^5}{120}U_i^{(5)} \ +\ \ldots\ \pm\ \mathcal{O}(h^7) \\
U_{i\pm2} &= U_i \ \pm\ 2hU'_i \ +\ \ldots\ \pm\ \frac{8h^3}{6}U_i^{(3)} \ +\ \ldots\ \pm\ \frac{32h^5}{120}U_i^{(5)} \ +\ \ldots\ \pm\ \mathcal{O}(h^7) \\
U'_{i\pm1} &= U'_i \ \pm\ \ldots\ +\ \frac{h^2}{2}U_i^{(3)} \ \pm \ldots\ +\ \frac{h^4}{24}U_i^{(5)} \ \pm\ \ldots\ +\ \mathcal{O}(h^6)
\end{aligned}$$

   Ensuite, on substitue ces développements dans la formule de Christian :

$$\begin{aligned}
(1+2\alpha)U'_i + 2\alpha\frac{h^2}{2}U_i^{(3)} + 2\alpha\frac{h^4}{24}U_i^{(5)} \quad &= \quad \frac{2\beta}{2h}\left(hU'_i + \frac{h^3}{6}U_i^{(3)} + \frac{h^5}{120}U_i^{(5)}\right) + \\
&\quad + \frac{2\gamma}{4h}\left(2hU'_i + \frac{8h^3}{6}U_i^{(3)} + \frac{32h^5}{120}U_i^{(5)}\right) \ +\ \mathcal{O}(h^6)
\end{aligned}$$

   En identifiant les termes en U'_i, $U_i^{(3)}$ et $U_i^{(5)}$, on obtient finalement :

$$\boxed{\begin{aligned} 2\alpha + 1 &= \beta + \gamma \\ 6\alpha &= \beta + 4\gamma \\ 2\alpha &= \beta + 16\gamma \end{aligned}}$$

   On obtient bien la précision pressentie, car le terme d'ordre six ne s'annule pas :-)

---

[3] Il s'agit de la valeur de l'exposant p dans le terme d'erreur $\mathcal{O}(h^p)$
Immédiatement, on peut observer que choisir $\alpha = \gamma = 0$ et $\beta = 1$ donne une différence finie centrée classique d'ordre deux.
Mais, cette réponse ne rendra heureux ni Christian, ni le correcteur...
C'est l'ordre le plus élevé possible qu'on souhaite avoir !

3. Calculer les valeurs optimales de α, β sachant que $\gamma = \frac{1}{9}$.

La solution unique du système des 3 équations : $\boxed{\alpha = \frac{1}{3}, \quad \beta = \frac{14}{9}, \quad \beta = \frac{1}{9}}$

Comme γ était fourni dans l'énoncé, on pouvait obtenir α et β à partir de deux équations. Inclure les trois valeurs dans la troisième équation permettait de vérifier si aucune erreur malencontreuse d'algèbre ne s'était pas glissée dans le calcul. La plupart des étudiants obtiennent ce résultat... tout en se trompant souvent sur l'ordre de précision de la première question.

4. Ecrire une fonction MATLAB

```
function [dU] =  compactDerivative(U,alpha,beta,gamma,h)
```

qui calcule le vecteur U_i' à partir des données U_i.
Le résultat sera fourni dans `dU` dont la taille sera celle de `U`.
Les coefficients α, β et γ ainsi que h sont donnés en argument d'entrée[4].

Une implémentation possible est :

```
function [dU] = compactDerivative(U,alpha,beta,gamma,h)
n = length(U);
A = sparse(n,n);
for i=2:n-1
        A(i,[i-1 i i+1]) = [alpha 1 alpha];
end
A(1,[n-1 1 2]) = [alpha 1 alpha];
A(n,[n-1 n 2]) = [alpha 1 alpha];
b = beta*(U([2:n 2]) - U([n-1 1:n-1]))/(2*h) ...
      + gamma*(U([3:n 2 3]) - U([n-2 n-1 1:n-2]))/(4*h);
dU = A\b;
end
```

Il faut <u>résoudre un système</u> pour obtenir la solution.
Ne pas l'observer fait perdre la totalité des points pour le programme.

Evidemment, il est essentiel d'utiliser une matrice creuse.
Quelques étudiants construisent deux autres matrices pour calculer le membre de droite : ce n'est pas formellement une erreur, mais c'est inutilement compliqué, mais cela n'a pas été pénalisé.

Il faut évidemment inclure les conditions de périodicité pour la matrice et le membre de droite. Noter que la valeur U_0 qui précède U_1 n'est pas U_n, mais U_{n-1}. La plupart des étudiants n'observent pas cette toute petite astuce, mais le correcteur ne leur en a pas tenu rigueur. Les plus astucieux observeront que le vecteur `U` doit avoir une taille minimale et être un vecteur colonne pour que le code fonctionne...

[4] Il est donc possible d'avoir un programme correct, même si on n'a pas obtenu les bonnes valeurs de α, β et γ.
Les données fournies sont périodiques : il n'est pas nécessaire de le tester.
Le code doit être simple et efficace.
Il n'est pas nécessaire d'écrire de commentaires :-)