Université de Rennes 2024-2025
Agrégation externe

Compléments sur les schémas numériques pour les équations différentielles

Miguel Rodrigues

On rappelle ici quelques éléments de théorie sur l'approximation des solutions des équations différentielles, parfois dites *équations différentielles ordinaires* par opposition aux équations aux dérivées partielles.
Ce que l'on discute ici peut être trouvé dans les livres classiques d'analyse numérique en général [Dem16, Fil13, Sch04]. Pour des traitements plus complets on pourra se reporter vers des livres spécialisés [CM92, CM97, HNW93, HW96].

1 Préliminaires

1.1 Position du problème

Soit $T>0$, $d\in\mathbf{N}^*$, $\Omega\subset\mathbf{R}^d$ ouvert, $u_0\in\mathbf{R}$ et $f:[0,T]\times\Omega\to\mathbf{R}^d$ continue et localement lipschitzienne par rapport à sa seconde [1] variable.
Supposons que le problème de Cauchy

$$(\forall t\in[0,T],\quad u'(t)=f(t,u(t))) \qquad\qquad u(0)=u_0\,, \tag{1.1}$$

possède une solution. L'hypothèse est ici que la solution est globale puisque, nécessairement, il existe une unique solution maximale d'après le théorème de Cauchy-Lipschitz.
Pour tout $N\in\mathbf{N}^*$, on se donne $(t_n^{(N)})_{0\leq n\leq N}$ une subdivision de $[0,T]$. On forcera que le pas maximal tend vers zéro lorsque le nombre de temps de discrétisation tend vers l'infini

$$h_N:=\max_{0\leq n\leq N-1}(t_{n+1}^{(N)}-t_n^{(N)})\stackrel{N\to\infty}{\longrightarrow}0\,.$$

Pour chaque $N\in\mathbf{N}^*$, on cherche à déterminer $(u_n^{(N)})_{0\leq n\leq N}$ approchant $(u(t_n^{(N)}))_{0\leq n\leq N}$. On dira que la méthode choisie pour déterminer cette approximation converge si

$$\max_{0\leq n\leq N}\|u_n^{(N)}-u(t_n^{(N)})\|\stackrel{N\to\infty}{\longrightarrow}0\,.$$

On dira que la méthode converge au moins à l'ordre $r>0$ s'il existe $C_r(f)$ tel que pour tout $N\in\mathbf{N}^*$

$$\max_{0\leq n\leq N}\|u_n^{(N)}-u(t_n^{(N)})\|\,\leq\,C_r(f)\,h_N^r\,.$$

On ne cherche à approcher u que sur une suite de temps discrets mais il est possible d'utiliser les valeurs $(u_n^{(N)})_{0\leq n\leq N}$ comme contraintes d'interpolation (polynomiale par morceaux) en $(t_n^{(N)})_{0\leq n\leq N}$ pour obtenir de u une approximation par une fonction. L'erreur d'une telle procédure cumule l'erreur sur les $(u(t_n^{(N)}))_{0\leq n\leq N}$ avec l'erreur du processus d'interpolation choisi.

1. Soulignons que l'ordre des variables n'est pas universel et qu'ici cela n'a de sens que parce que nous avons d'abord prescrit cet ordre.

Dans la suite on se focalisera sur le cas où la valeur $u_{n+1}^{(N)}$ est calculée uniquement à partir de $t_n^{(N)}$, $t_{n+1}^{(N)}$ et $u_n^{(N)}$. On pourrait cependant imaginer utiliser plus de valeurs précédentes.

Au moins implicitement, le choix d'un problème approché est basé sur la reformulation suivante du problème de Cauchy : pour tout $N \in \mathbf{N}^*$, u résout (1.1) si et seulement si $u(t_0^{(N)}) = u_0$ et pour tout $0 \leq n \leq N-1$,

$$(\forall t \in [t_n^{(N)}, t_{n+1}^{(N)}], \quad u'(t) = f(t, u(t))) \qquad\qquad u(t_n^{(N)}) = u(t_n^{(N)})\,,$$

c'est-à-dire

$$\forall t \in [t_n^{(N)}, t_{n+1}^{(N)}], \quad u(t) = u(t_n^{(N)}) + \int_{t_n^{(N)}}^{t} f(s, u(s)))\,\mathrm{d}s\,.$$

En particulier, pour tout $0 \leq n \leq N-1$,

$$u(t_{n+1}^{(N)}) = u(t_n^{(N)}) + \int_{t_n^{(N)}}^{t_{n+1}^{(N)}} f(s, u(s)))\,\mathrm{d}s\,. \tag{1.2}$$

On notera que le problème contient les problèmes d'intégration numérique qui correspondent aux cas où f ne dépend pas de sa seconde variable. De cette remarque on peut déjà déduire que l'on aura également besoin de régularité par rapport à la première variable de f.

1.2 Stabilité par rapport aux perturbations dans l'équation

Les méthodes étudiées ici sont construites en perturbant (1.2) pour obtenir un moyen de passer de $u_n^{(N)}$ à $u_{n+1}^{(N)}$. À titre d'échauffement mais aussi pour comprendre ce que l'on pourra espérer de mieux, commençons par examiner l'impact sur les solutions des perturbations sur l'équation différentielle.

Proposition 1 *Soit* $0 < \delta < d(u([0,T]), \partial\Omega)$ *et* L *une constante de Lipschitz*[2] *en sa seconde variable pour la restriction de* f *à* $[0,T] \times (u([0,T]) + \overline{B}(0,\delta))$. *Pour tout* $v_0 \in \mathbf{R}^d$ *et* $\varepsilon : [0,T] \to \mathbf{R}^d$ *continue tels que*

$$e^{L\,T}\left(\|u_0 - v_0\| + \int_0^T \|\varepsilon(s)\|\,ds\right) < \delta$$

alors le problème de Cauchy

$$(\forall t \in [0,T], \quad v'(t) = f(t, v(t)) + \varepsilon(t)) \qquad\qquad v(0) = v_0\,,$$

possède une unique solution et pour tout $0 \leq t \leq T$,

$$\|u(t) - v(t)\| \leq e^{L\,t}\left(\|u_0 - v_0\| + \int_0^t e^{-L\,s}\|\varepsilon(s)\|\,ds\right)\,.$$

On notera que pour contrôler ponctuellement l'erreur de la solution on a besoin de contrôler une intégrale de l'erreur dans l'équation, dit autrement on contrôle la solution dans L^∞ à partir d'un contrôle L^1 de l'équation. C'est naturel puisque l'on intègre l'équation pour trouver la solution.

La constante de Lipschitz associée à la proposition est $\mathrm{e}^{L\,T}$ où L est une constante de Lipschitz pour f. La démonstration utilise le principe de continuité.

2. On utilise ici que la restriction à un compact d'une fonction localement lipschitzienne est lipschitzienne. Le démontrer est un bon exercice.

Démonstration. Notons I l'ensemble des temps $0 \leq t \leq T$, tels que la solution maximale v soit définie sur $[0, t]$ et pour tout $0 \leq s \leq t$

$$\|u(s) - v(s)\| \leq \mathrm{e}^{L\,s}\left(\|u_0 - v_0\| + \int_0^s \mathrm{e}^{-L\,\tau}\|\varepsilon(\tau)\|\mathrm{d}\tau\right).$$

Par connexité de $[0, T]$, puisque $0 \in I$, il suffit de montrer que I est ouvert et fermé dans $[0, T]$.
Supposons que $(t_n)_{n\in\mathbf{N}} \in I^{\mathbf{N}}$ converge vers $t \in [0, T]$. La définition de I implique que $(v(t_n))_{n\in\mathbf{N}}$ est à valeurs dans le compact $u([0, T]) + \overline{B}(0, \delta)$. Cela implique que t n'est pas le temps maximal d'existence de v. Ainsi v est bien définie sur $[0, t]$ et en passant à la limite on déduit que $t \in I$. Ainsi I est fermé.
Soit $t \in I$. Alors il est clair que tout $0 \leq s \leq t$ appartient I. Dans le cas où $t = T$ cela suffit à montrer que I contient un voisinage de t dans $[0, T]$. Supposons $t < T$. Alors d'après le théorème de Cauchy-Lipschitz, il existe $t < t' \leq T$ tel que v soit défini sur $[0, t']$. Comme $t \in I$, on déduit que $\|u(t) - v(t)\| < \delta$ et donc, par continuité, que quitte à réduire t' on peut supposer que pour tout $0 \leq s \leq t'$, $\|u(s) - v(s)\| < \delta$. Ainsi pour tout $0 \leq s \leq t'$ on peut alors déduire de

$$v(s) - u(s) \;=\; v_0 - u_0 + \int_0^s (f(\tau, v(\tau)) - f(\tau, u(\tau)))\,\mathrm{d}\tau + \int_0^s \varepsilon(\tau)\,\mathrm{d}\tau$$

que

$$\|v(s) - u(s)\| \;\leq\; \|v_0 - u_0\| + L\int_0^s \|v(\tau) - u(\tau)\|\,\mathrm{d}\tau + \int_0^s \|\varepsilon(\tau)\|\,\mathrm{d}\tau\,.$$

Du lemme de Grönwall on conclut alors que pour tout $0 \leq s \leq t'$

$$\|u(s) - v(s)\| \leq \mathrm{e}^{L\,s}\left(\|u_0 - v_0\| + \int_0^s \mathrm{e}^{-L\,\tau}\|\varepsilon(\tau)\|\mathrm{d}\tau\right).$$

Ainsi $[0, t']$ est un voisinage de t inclus dans I. I est donc également ouvert dans $[0, T]$.
D'où $I = [0, T]$. ■

Comme souvent dans les problèmes d'équations différentielles si l'on suppose que $\Omega = \mathbf{R}^d$ et que f est globalement lipschitzienne en sa seconde variable la démonstration précédente est significativement simplifiée parce que l'on peut couper les parties argument de continuité/solutions maximales. Cette hypothèse réduit significativement les f que l'on peut traiter mais permet de ne garder que la partie de la démonstration spécifique. Il ne faut pas hésiter à faire ce type d'hypothèse pour se simplifier la démonstration ou pour mieux souligner le cœur de l'argument.

2 Le schéma d'Euler explicite

2.1 Présentation

Le schéma d'Euler explicite s'écrit : pour tout $0 \leq n \leq N - 1$,

$$\frac{u_{n+1}^{(N)} - u_n^{(N)}}{t_{n+1}^{(N)} - t_n^{(N)}} \;=\; f(t_n^{(N)}, u_n^{(N)})$$

ou encore

$$u_{n+1}^{(N)} = u_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\,f(t_n^{(N)}, u_n^{(N)})\,. \tag{2.1}$$

Quand Ω n'est pas $\mathbf{R}^d$ tout entier il est possible que l'on ne puisse pas résoudre (2.1) jusqu'en $n = N-1$, ce qui est une contre-partie discrète à l'explosion en temps fini.
La première forme souligne que le schéma d'Euler explicite peut être vu comme découlant de l'approximation de la dérivée par un taux d'accroissement à droite alors que la seconde forme souligne qu'il peut être aussi vu comme découlant de la méthode des rectangles à gauche appliquée à la formulation intégrale (1.2).
Il est plutôt plus facile de généraliser le second point de vue parce que l'on connaît plusieurs méthodes de quadrature. On notera que si l'on change quoi que ce soit à l'approximation du schéma d'Euler explicite on se retrouve à évaluer u en des temps pour lesquels on n'a pas encore calculé d'approximation... On y reviendra.
On ne peut pas espérer que la méthode converge plus qu'à l'ordre 1 pour des f généraux puisque c'est déjà le cas pour la méthode des rectangles. Le but est de montrer que c'est bien le cas pour des f assez réguliers.
Nous allons séparer l'analyse de convergence en deux parties :

1. l'analyse de *consistance*, c'est-à-dire ici l'analyse d'à quel point la solution de l'équation différentielle u fournit une bonne solution approchée $(u(t_n^{(N)}))_{0\leq n\leq N}$ du problème discret (2.1) ;
2. l'analyse de *stabilité*, c'est-à-dire ici l'analyse de comment une erreur dans le schéma (2.1) se répercute en erreur sur sa solution.

Si l'on fait des choix compatibles de normes pour la stabilité et la consistance, conclure la convergence est presque immédiat. Insistons sur le fait que dans l'analyse de consistance il s'agit de vérifier que notre problème discret approche bien le problème continu qu'il doit approcher dans la limite de petit pas de discrétisation, alors que dans l'analyse de stabilité il faut vérifier que la famille de problèmes discrets est une bonne famille de problèmes *uniformément par rapport au pas de discrétisation.*

2.2 Analyse de stabilité

Commençons par l'analyse de stabilité, cela nous guidera dans le choix des normes pour l'analyse de consistance. Puisque la stabilité doit être uniforme en la discrétisation elle doit être compatible à la limite avec le résultat, au niveau continu, de la proposition 1.
Le lemme suivant est ainsi une contre-partie discrète de la proposition 1 (sans sa partie argument de continuité/solutions maximales).

Lemme 2 *Soit K un compact de $\mathbf{R}^d$ inclus dans Ω et L_K la constante de Lipschitz en sa seconde variable pour la restriction de f à $[0,T]\times K$. Pour tout $N\in\mathbf{N}^*$, si pour un certain entier $0\leq N'\leq N$, $(u_n^{(N)})_{0\leq n\leq N'}$, $(v_n^{(N)})_{0\leq n\leq N'}$ et $(\varepsilon_n^{(N)})_{0\leq n\leq N'}$ vérifient que pour tout $0\leq n\leq N'-1$, $u_n^{(N)}\in K$, $v_n^{(N)}\in K$ et*

$$\begin{aligned}\frac{u_{n+1}^{(N)}-u_n^{(N)}}{t_{n+1}^{(N)}-t_n^{(N)}} &= f(t_n^{(N)},u_n^{(N)})\\ \frac{v_{n+1}^{(N)}-v_n^{(N)}}{t_{n+1}^{(N)}-t_n^{(N)}} &= f(t_n^{(N)},v_n^{(N)})+\varepsilon_n^{(N)}\end{aligned}$$

ou encore

$$
\begin{aligned}
u_{n+1}^{(N)} &= u_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, f(t_n^{(N)}, u_n^{(N)})\\
v_{n+1}^{(N)} &= v_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, f(t_n^{(N)}, v_n^{(N)}) + (t_{n+1}^{(N)} - t_n^{(N)})\, \varepsilon_n^{(N)}
\end{aligned}
$$

alors pour tout $0 \leq n \leq N'$

$$
\|u_n^{(N)} - v_n^{(N)}\| \leq e^{L_K\, T} \left(\|u_0^{(N)} - v_0^{(N)}\| + \sum_{\ell=0}^{n-1} (t_{\ell+1}^{(N)} - t_\ell^{(N)})\, \|\varepsilon_\ell^{(N)}\| \right).
$$

Démonstration. On obtient immédiatement que, pour tout $0 \leq n \leq N' - 1$,

$$
\|u_{n+1}^{(N)} - v_{n+1}^{(N)}\| \leq L_K\, (t_{n+1}^{(N)} - t_n^{(N)})\, \|u_n^{(N)} - v_n^{(N)}\| + (t_{n+1}^{(N)} - t_n^{(N)})\, \|\varepsilon_n^{(N)}\|.
$$

On déduit alors par récurrence que, pour tout $0 \leq n \leq N$,

$$
\begin{aligned}
\|u_n^{(N)} - v_n^{(N)}\| \leq & \left(\prod_{\ell=0}^{n-1} (L_K\, (t_{\ell+1}^{(N)} - t_\ell^{(N)})) \right) \|u_0^{(N)} - v_0^{(N)}\| \\
& + \sum_{\ell=0}^{n-1} \left(\prod_{j=\ell}^{n-1} (L_K\, (t_{j+1}^{(N)} - t_j^{(N)})) \right) (t_{\ell+1}^{(N)} - t_\ell^{(N)})\, \|\varepsilon_\ell^{(N)}\|.
\end{aligned}
$$

On conclut en invoquant que, pour tout $x \geq 0$, l'on a $x \leq \mathrm{e}^x$ de sorte que, pour tout $0 \leq n \leq N$,

$$
\prod_{\ell=0}^{n-1} (L_K\, (t_{\ell+1}^{(N)} - t_\ell^{(N)})) \leq \prod_{\ell=0}^{n-1} \mathrm{e}^{L_K\, (t_{\ell+1}^{(N)} - t_\ell^{(N)})} = \exp\left(\sum_{\ell=0}^{n-1} (L_K\, (t_{\ell+1}^{(N)} - t_\ell^{(N)})) \right) = \exp\left(L_K (t_n^{(N)} - t_0^{(N)} \right)
$$

et de même pour tout $0 \leq \ell \leq n$

$$
\prod_{j=\ell}^{n-1} (L_K\, (t_{j+1}^{(N)} - t_j^{(N)})) \leq \mathrm{e}^{L_K (t_n^{(N)} - t_\ell^{(N)})}.
$$

Ainsi, pour tout $0 \leq n \leq N$,

$$
\|u_n^{(N)} - v_n^{(N)}\| \leq \mathrm{e}^{L_K\, t_n^{(N)}} \left(\|u_0^{(N)} - v_0^{(N)}\| + \sum_{\ell=0}^{n-1} (t_{\ell+1}^{(N)} - t_\ell^{(N)})\, \|\varepsilon_\ell^{(N)}\|\, \mathrm{e}^{-L_K\, t_\ell^{(N)}} \right).
$$

D'où le résultat. ∎

Si l'on souhaite se simplifier les notations on peut raisonnablement se limiter au cas où les pas de temps sont équidistants.

2.3 Analyse de consistance

On rappelle que l'on note u la solution de (1.1). Pour la méthode d'Euler explicite, on appelle erreur de consistance locale les

$$
\varepsilon_n^{(N)} := \frac{u(t_{n+1}^{(N)}) - u(t_n^{(N)})}{t_{n+1}^{(N)} - t_n^{(N)}} \; - \; f(t_n^{(N)}, u(t_n^{(N)})),
$$

et erreur de consistance globale les

$$
\mathcal{E}^{(N)} := \sum_{\ell=0}^{N-1} (t_{\ell+1}^{(N)} - t_\ell^{(N)})\, \|\varepsilon_\ell^{(N)}\|.
$$

Ces définitions peuvent être généralisées à d'autres schémas.
On dit que la méthode numérique est consistante si

$$\mathcal{E}^{(N)} \overset{N\to\infty}{\longrightarrow} 0$$

et qu'elle est consistante au moins à l'ordre $r > 0$ s'il existe $C_r(f)$ tel que pour tout $N \in \mathbf{N}^*$

$$\mathcal{E}^{(N)} \leq C_r(f)\, h_N^r .$$

Le lemme suivant montre que la méthode d'Euler explicite est consistante au moins à l'ordre 1.

Lemme 3 *Supposons f de classe $\mathcal{C}^1$ de sorte que u soit de classe $\mathcal{C}^2$. Alors, pour tout $N \in \mathbf{N}^*$,*

$$\mathcal{E}^{(N)} \leq \frac{1}{2} T \left(\max_{[0,T]} |u''|\right) h_N .$$

Démonstration. Notons d'abord que

$$\varepsilon_n^{(N)} = \int_0^1 \left(u'(t_n^{(N)} + s\,(t_{n+1}^{(N)} - t_n^{(N)})) - u'(t_n^{(N)})\right) \mathrm{d}s$$

de sorte que

$$\|\varepsilon_n^{(N)}\| \leq (t_{n+1}^{(N)} - t_n^{(N)}) \max_{[0,T]} |u''| \int_0^1 s\, \mathrm{d}s = \frac{1}{2} (t_{n+1}^{(N)} - t_n^{(N)}) \max_{[0,T]} |u''| .$$

D'où

$$\mathcal{E}^{(N)} \leq \frac{1}{2} \max_{[0,T]} |u''| \sum_{\ell=0}^{N-1} (t_{n+1}^{(N)} - t_n^{(N)})^2 \leq \frac{1}{2} \max_{[0,T]} |u''|\, h_N \sum_{\ell=0}^{N-1} (t_{n+1}^{(N)} - t_n^{(N)}) = \frac{1}{2} T \max_{[0,T]} |u''|\, h_N .$$

■

2.4 Analyse de convergence

Il suffit maintenant de combiner la consistance et la stabilité en complétant par un équivalent discret de l'argument de continuité. Ce dernier est plutôt plus simple ; il repose sur le principe de récurrence plutôt que sur la connexité.

Théorème 4 *Supposons f $\mathcal{C}^1$. Soit $0 < \delta < d(u([0,T]), \partial\Omega)$ et L une constante de Lipschitz en sa seconde variable pour la restriction de f à $[0,T] \times (u([0,T]) + \overline{B}(0,\delta))$. Pour tout pas de temps h_N suffisamment petit pour assurer*

$$\frac{1}{2} T\, e^{LT} \max_{[0,T]} |u''|\, h_N < \delta ,$$

la solution $(u_n^{(N)})_{0 \leq n \leq N}$ du schéma d'Euler explicite est bien définie et vérifie, pour tout $0 \leq n \leq N$,

$$\|u(t_n^{(N)}) - u_n^{(N)}\| \leq \frac{1}{2} T\, e^{LT} \max_{[0,T]} |u''|\, h_N .$$

Démonstration. Appliquons le lemme 2 avec $K = u([0,T]) + \overline{B}(0,\delta)$ et $v_n^{(N)} := u(t_n^{(N)})$. On déduit alors grâce au lemme 3 par récurrence (finie) sur n que, pour tout $0 \leq n \leq N$, $u_n^{(N)} \in K$ et que l'estimation annoncée est vraie au rang n. ■

3 D'autres schémas

Pour analyser d'autres schémas où la valeur $u_{n+1}^{(N)}$ est calculée uniquement à partir de $t_n^{(N)}$, $t_{n+1}^{(N)}$ et $u_n^{(N)}$, il est naturel d'écrire la relation de récurrence sous la forme

$$u_{n+1}^{(N)} = u_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, \Phi_f(t_n^{(N)}, u_n^{(N)}, t_{n+1}^{(N)} - t_n^{(N)})\,.$$

Le cas du schéma d'Euler explicite correspond au cas

$$\Phi_f(t,u,h) = f(t,u)\,.$$

On définirait alors les erreurs de consistance locale par

$$\varepsilon_n^{(N)} := \frac{u(t_{n+1}^{(N)}) - u(t_n^{(N)})}{t_{n+1}^{(N)} - t_n^{(N)}} \; - \; \Phi_f(t_n^{(N)}, u(t_n^{(N)}), t_{n+1}^{(N)} - t_n^{(N)})\,.$$

Nous n'allons pas mener l'analyse d'autres schémas, mais donnons quelques idées de ce qu'il faudrait faire. On peut généraliser la démonstration de la stabilité (discrète uniformément par rapport à la discrétisation) pourvu que Φ_f soit localement lipschitzienne par rapport à sa deuxième variable (ce qui n'est pas une contrainte forte). Pour examiner l'ordre de convergence il suffit de calculer le développement lorsque $h \to 0$ de

$$\frac{u(t+h) - u(t)}{h} - \Phi_f(t, u(t), h)$$

lorsque u résout (1.1). Pour réaliser cette dernière tâche il est utile d'exprimer les dérivées de u au temps t comme des fonctions de $(t, u(t))$, ce que l'on réalise par récurrence : si la fonction f_p est telle que pour tout t, $u^{(p)}(t) = f_p(t, u(t))$, alors $f_{p+1}(t,u) := \partial_t f_p(t,u) + \mathrm{d}_u f_p(t,u)(f(t,u))$ définit une fonction f_{p+1} vérifiant la propriété au rang suivant.
On obtient par exemple ainsi que la méthode est au moins d'ordre 1 pourvu que pour tout (t,v)

$$\Phi_f(t,v,0) = f(t,v)$$

et qu'elle est au moins d'ordre 2 pourvu que pour tout (t,v)

$$\Phi_f(t,v,0) = f(t,v)\,, \qquad\qquad \partial_h \Phi_f(t,v,0) = \frac{1}{2}\left(\partial_t f(t,v) + \mathrm{d}_u f(t,v)(f(t,v))\right)\,.$$

3.1 Quelques (autres) schémas de Runge-Kutta

La raison la plus courante pour utiliser une autre méthode que la méthode d'Euler explicite est la volonté d'avoir une méthode d'ordre plus élevé. La méthode la plus couramment utilisée (en boîte noire) est une méthode d'ordre 4, appelée méthode de Runge-Kutta d'ordre 4 et qui est une généralisation de la méthode de quadrature de Simpson.
Contentons-nous de montrer comment obtenir des méthodes d'ordre 2. Commençons par montrer comment généraliser la méthode du point milieu. Le problème est qu'avant d'appliquer la formule du point milieu il faut avoir une approximation de la valeur de la solution après un demi-pas de temps. À partir d'une approximation $u_n^{(N)}$ au temps $t_n^{(N)}$, on peut calculer une approximation $u_{n+\frac12}^{(N)}$ au temps $t_{n+\frac12}^{(N)} := (t_n^{(N)} + t_{n+1}^{(N)})/2$ par

$$u_{n+\frac12}^{(N)} = u_n^{(N)} + \tfrac12\,(t_{n+1}^{(N)} - t_n^{(N)})\, f(t_n^{(N)}, u_n^{(N)})$$

(schéma d'Euler explicite, méthode des rectangles à gauche) puis poser

$$u_{n+1}^{(N)} = u_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, f(t_{n+\frac{1}{2}}^{(N)}, u_{n+\frac{1}{2}}^{(N)})\,.$$

Cela correspond au choix

$$\Phi_f(t,u,h) = f(t + \tfrac{1}{2}h, u + \tfrac{1}{2}hf(t,u))\,.$$

Montrons maintenant comment généraliser la méthode des trapèzes. À partir d'une approximation $u_n^{(N)}$ au temps $t_n^{(N)}$, on peut calculer une première approximation $\tilde{u}_{n+1}^{(N)}$ au temps $t_{n+1}^{(N)}$ par

$$\tilde{u}_{n+1}^{(N)} = u_n^{(N)} + \,(t_{n+1}^{(N)} - t_n^{(N)})\, f(t_n^{(N)}, u_n^{(N)})$$

(schéma d'Euler explicite, méthode des rectangles à gauche) puis poser

$$u_{n+1}^{(N)} = u_n^{(N)} + \frac{1}{2}(t_{n+1}^{(N)} - t_n^{(N)})\, \left(f(t_n^{(N)}, u_n^{(N)}) + f(t_{n+1}^{(N)}, \tilde{u}_{n+1}^{(N)})\right)\,.$$

Cela correspond au choix

$$\Phi_f(t,u,h) = \frac{1}{2}\left(f(t,u) + f(t+h, u + hf(t,u))\right)\,.$$

Les calculs de consistance d'ordre élevé posent vite des défis de calcul et de représentation de nature essentiellement algébriques.

3.2 Schéma d'Euler implicite

Jusqu'ici nous nous sommes imposés la contrainte que la valeur au temps suivant est donnée directement par une formule, et pas par la résolution (a priori seulement approchée) d'une équation. Quand ce n'est pas le cas on parle de méthode implicite.

Évidemment il faut une excellente raison pour choisir un schéma implicite et compter sur le fait que l'équation à résoudre n'est pas trop compliquée.

La principale raison pour le choix d'une méthode implicite est d'améliorer les propriétés de stabilité du schéma. C'est nécessaire quand les constantes de Lipschitz de f sont grandes et surestiment la séparation des trajectoires. On parle de problème *raide*.

Illustrons-le par un cas trivial. Considérons le cas où $\Omega = \mathbf{R}$ et une famille de champs de vecteurs est donnée par $f_\epsilon(t,u) = -\frac{1}{\epsilon}u$. Alors localement uniformément sur $]0,T]$, la solution u_ϵ du problème de Cauchy converge vers la fonction nulle lorsque ϵ tend vers zéro par valeurs positives. On aimerait pouvoir calculer une approximation d'une fonction presque nulle à peu de frais. Si l'on utilise le schéma d'Euler explicite avec des pas équidistants de longueur $h_N = T/N$, l'approximation discrète est donnée par

$$u_n^{(N)} = \left(1 - \frac{1}{\epsilon}h_N\right)^n\, u_0^{(N)}$$

de sorte que

$$u_N^{(N)} = \left(1 - \frac{1}{\epsilon}h_N\right)^{\frac{T}{h_N}}\, u_0^{(N)}\,.$$

Pour empêcher l'explosion lorsque $\epsilon \to 0$, il faut contraindre N par $h_N \leq \epsilon$ et donc $N \geq T/\epsilon$, ce qui impose un nombre de pas gigantesque.

Une généralisation naturelle de la méthode des rectangles à droite consiste à calculer $u_{n+1}^{(N)}$ en résolvant

$$u_{n+1}^{(N)} = u_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, f(t_{n+1}^{(N)}, u_{n+1}^{(N)})$$

ce qui correspond à un $\Phi_f(t,u)$ défini implicitement par

$$u + h\,\Phi_f(t,u) - h\,f(t+h, u + h\,\Phi_f(t,u)) = u$$

Il s'agit de la méthode d'Euler implicite.
Sur l'exemple précédent, avec le schéma d'Euler implicite avec des pas équidistants de longueur $h_N = T/N$, l'approximation discrète est donnée par

$$u_n^{(N)} = \frac{1}{\left(1 + \frac{1}{\epsilon} h_N\right)^n}\, u_0^{(N)}$$

ce qui donne le bon comportement lorsque $\epsilon \to 0^+$ sans contrainte sur N.
Une autre manière de traiter les problèmes raides consistent à adapter les pas de temps aux valeurs que l'on calcule pour prendre des pas de temps petits là où la solution varie vite et des pas de temps grands quand elle varie peu. On parle de méthodes *à pas adaptatifs*. Sur l'exemple ci-dessus il suffit de prendre des pas de temps de taille ϵ seulement sur les premiers pas de temps. Soulignons que l'on ne sait pas a priori où la solution varie vite et il faut donc que le schéma le détecte lui-même au cours du calcul à partir des approximations. Beaucoup de méthodes préprogrammées le font.
Donnons un exemple moins trivial mais plus difficile à analyser d'utilisation de schémas implicites. Les solutions $u_\epsilon = (x_\epsilon, y_\epsilon)$ du système

$$\begin{aligned} x_\epsilon' &= y_\epsilon\,, \\ y_\epsilon' &= -\frac{1}{\epsilon}\,(y_\epsilon - g(x_\epsilon))\,, \end{aligned}$$

sont telles que x_ϵ converge lorsque $\epsilon \to 0^+$ vers une solution de $x' = g(x)$. Pour démontrer l'assertion précédente on peut commencer par noter que le système implique

$$(x_\epsilon + \epsilon\, y_\epsilon)' = g(x_\epsilon)\,.$$

Un schéma implicite reproduisant cette propriété sans contrainte sur le pas est

$$\begin{aligned} x_{n+1}^{(N)} &= x_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, y_{n+1}^{(N)}\,, \\ y_{n+1}^{(N)} &= y_n^{(N)} - \frac{1}{\epsilon}\,(t_{n+1}^{(N)} - t_n^{(N)})\,\left(y_{n+1}^{(N)} - g(x_n^{(N)})\right)\,. \end{aligned}$$

Pour s'en convaincre on peut commencer par noter que le schéma implique

$$x_{n+1}^{(N)} + \epsilon y_{n+1}^{(N)} = x_n^{(N)} + \epsilon y_n^{(N)} + (t_{n+1}^{(N)} - t_n^{(N)})\, g(x_n^{(N)})\,.$$

Références

[CM92] M. Crouzeix and A.-L. Mignot. *Analyse numérique des équations différentielles.* Masson, 1992.

[CM97] M. Crouzeix and A.-L. Mignot. *Exercices d'analyse numérique des équations différentielles.* Masson, 1997.

[Dem16] J.-P. Demailly. *Analyse numérique et équations différentielles.* ÉDP Sciences, 2016.

[Fil13] F. Filbet. *Analyse numérique - Algorithme et étude mathématique.* Dunod, 2013.

[HNW93] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, second edition, 1993. Nonstiff problems.

[HW96] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. Stiff and differential-algebraic problems.

[Sch04] M. Schatzman. *Analyse numérique*. Dunod, 2004.