

---

## Méthodes numériques pour les EDO.

---

### 3.1 Motivations et principes généraux.

Nous avons vu dans la deuxième partie du cours que, sous des conditions assez générales, le problème de Cauchy pour une équation différentielle admet une unique solution. Mais dans de nombreux cas, même si la solution existe, il est impossible de la calculer analytiquement.

On cherche alors à construire une **solution approchée** qui soit effectivement calculable, par exemple à l'aide d'un ordinateur. Cela implique que la méthode pour calculer cette solution approchée ne fasse intervenir qu'un nombre fini d'inconnues (alors que la vraie solution, la **solution exacte**, qui est une fonction, est définie par un nombre infini et non dénombrable de valeurs).

Nous allons construire cette solution approchée par une **méthode numérique**, que l'on appelle aussi souvent de **schéma numérique**. Nous allons présenter plusieurs méthodes (schémas) numériques pour calculer une solution approchée d'une EDO, puis nous allons étudier la **convergence** de ces méthodes vers la solution exacte du problème de Cauchy. Nous allons voir que la convergence d'une méthode porte sur deux propriétés essentielles, que l'on étudiera : la **consistance** et la **stabilité**. On abordera aussi la notion d'**ordre de convergence**, liée à la vitesse à laquelle la méthode converge.

## 3.2 La notion de schéma, ou méthode, numérique. Définitions et exemples.

Considérons le problème de Cauchy

$$(PC) \quad \begin{cases} y' = f(t, y), \\ y(t_0) = y_0, \end{cases}$$

où  $f : I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ , avec  $I$  intervalle ouvert de  $\mathbb{R}$ , est une fonction continue et localement Lipschitzienne par rapport à sa deuxième variable et où  $(t_0, y_0) \in I \times \mathbb{R}^n$  est donné.

*Pour alléger les notations, nous considérons ici le cas d'une fonction définie dans un ensemble de la forme  $I \times U$ , avec  $I$  intervalle ouvert de  $\mathbb{R}$  et  $U = \mathbb{R}^n$  ; toutes les notions et résultats qu'on verra dans ce chapitre restent valables dans le cas où  $U$  est un ouvert strictement contenu dans  $\mathbb{R}^n$ .*

Dans ces conditions, le théorème de Cauchy Lipschitz et ses corollaires impliquent que le problème de Cauchy (PC) admet une unique solution maximale  $y : J \longrightarrow \mathbb{R}^n$  définie dans un intervalle  $J \subseteq I$  tel que  $t_0 \in J$ .

Notre objectif est d'approcher la solution  $y$  de (PC) dans un intervalle de la forme  $[t_0, t_0 + T] \subseteq J$ , avec  $T > 0$ , par une **solution discrète** (c'est-à-dire une fonction dont le domaine de définition est un ensemble discret).

Pour cela on procède comme suit.

- i) On commence par discrétiser (c'est-à-dire remplacer du continu par du discret) l'intervalle  $[t_0, t_0 + T]$ . Pour ce faire on se donne  $N \in \mathbb{N}$  et on définit une sub-division de  $[t_0, t_0 + T]$  en  $N$  sous-intervalles définis par les  $N + 1$  points, appelées points de la discrétisation,

$$t_0 < t_1 < \dots < t_N = t_0 + T.$$

*On appelle aussi les points  $t_n$  les instants de temps, car dans le cas de nombreuses EDO modélisant des phénomènes physiques, la variable  $t$  représente le temps.*

Le point  $t_0$  est appelé l'**instant initial** et le point  $t_N$  l'**instant final**. On appelle aussi :

$h_n := t_{n+1} - t_n$  le **pas de temps** ou le **pas de discrétisation** entre  $t_n$  et  $t_{n+1}$ , pour  $n = 0, \dots, N - 1$  ;

$h := \max_{n=0, \dots, N-1} h_n$  le **pas de temps** ou **pas de la discrétisation**.

La discrétisation est dite **uniforme** si  $h_n = h$ , pour tout  $n = 0, \dots, N - 1$ . Dans ce cas la pas  $h$  de la discrétisation est défini par

$$h = \frac{T}{N}$$

et les  $N + 1$  points de discrétisation sont définis par

$$t_n = t_0 + nh, \quad n = 0, \dots, N.$$

- ii) On construit  $N + 1$  valeurs  $y^0, \dots, y^N$  qui approchent la solution exacte  $y$  de  $(PC)$ , aux points  $t_0, \dots, t_N$ , autrement dit on construit des valeurs  $y^0, \dots, y^N$  tels que

$$y^n \approx y(t_n), \text{ pour } n = 0, \dots, N.$$

Une **méthode numérique** ou un **schéma numérique** est donc la donnée d'une telle construction. Plus précisément, c'est la donnée d'une suite  $(y^0, \dots, y^N)_{N \in \mathbb{N}}$ , indexée par  $N \in \mathbb{N}$ , de valeurs approchant les valeurs exactes  $(y(t_0), \dots, y(t_N))$  de la solution  $y$  de  $(PC)$  aux points de la discrétisation  $(t_0, \dots, t_N)$ . On appelle souvent la suite  $(y^0, \dots, y^N)_{N \in \mathbb{N}}$  de **solution approchée** ou **solution numérique** et, pour  $N \in \mathbb{N}$  donnée, le terme d'ordre  $N$  de cette suite solution approchée (ou solution numérique) associée à une discrétisation de l'intervalle  $[t_0, t_0 + T]$  de  $N + 1$  points ou de pas  $h = \max_{0, \dots, N} h_n$ .

Comme la valeur de la solution exacte à l'instant initial  $y(t_0) = y_0$  est connue, on prendra  $y^0 = y_0$ .

La dépendance de la suite  $(y^0, \dots, y^N)$  par rapport à  $N$  n'est pas toujours spécifiée mais il faut la garder à l'esprit, ainsi que la dépendance de la suite  $(t_0, \dots, t_N)$  définissant les points de la discrétisation de l'intervalle  $[t_0, t_0 + T]$ .

*Lorsque  $N$  augment, le nombre de points de la discrétisation augmente et on s'attend à ce que les valeurs approchées  $y^0, \dots, y^N$  approchent de manière de plus en plus précise les valeurs exactes  $y(t_0), \dots, y(t_N)$ .*

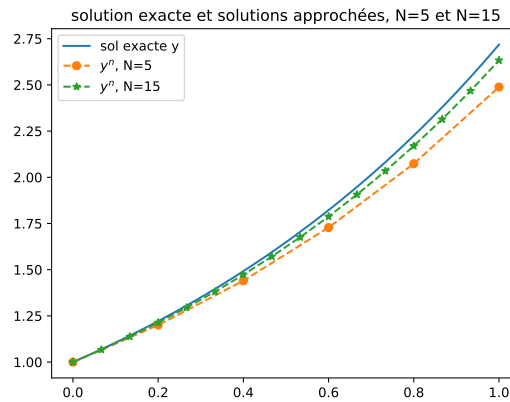
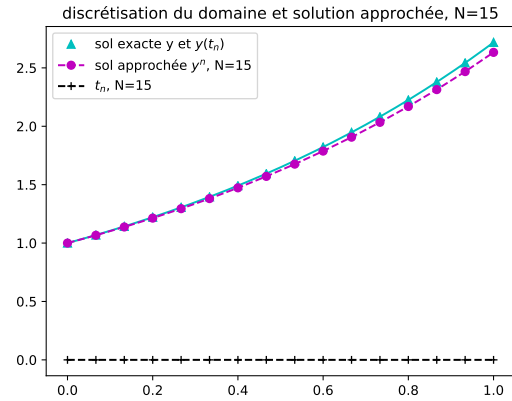
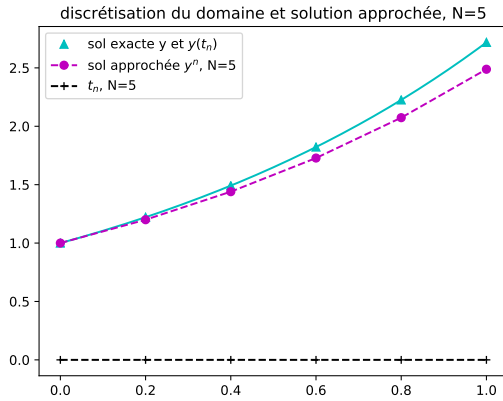
Nous allons nous restreindre à des discrétisations uniformes de l'intervalle  $[t_0, t_0 + T]$  (le cas d'une discrétisation à pas  $h_n$  variable est souvent utilisé lorsque l'on veut adapter le pas de la discrétisation pour améliorer la précision du schéma). Dans ce cas, comme  $h = \frac{T}{N}$ , on peut de manière équivalente voir la suite approchée  $(y^0, \dots, y^N)$  comme une suite indexée par  $N \in \mathbb{N}$  ou comme une suite indexée par  $h \in ]0, 1]$ ; on a choisi l'intervalle  $]0, 1]$  pour  $h$  mais on aurait pu choisir n'importe quel autre intervalle de la forme  $]0, h_{\max}]$ . Ce qui nous intéressera sera le comportement de la solution approchée  $(y^0, \dots, y^N)$  lorsque  $N$  tend vers l'infini (*i.e.*, lorsque le nombre de points de la discrétisation tend vers l'infini, ce qui correspond à ce que l'ensemble discret qui représente l'intervalle  $[t_0, t_0 + T]$  tend vers cet intervalle), ou de manière équivalente, lorsque  $h$  tend vers 0 (*i.e.* lorsque la distance entre deux points consécutifs de la discrétisation tend vers 0).

Dans la figure suivante on illustre les notions que l'on vient de présenter. On y représente la fonction exponentielle  $y(t) = e^t$ , qui est la solution du problème de Cauchy

$$\begin{cases} y' = y, \\ y(0) = 1, \end{cases}$$

sur l'intervalle  $[0, 1]$ , ainsi que les solutions approchées  $(y^0, \dots, y^N)$  associées à des discrétisations de l'intervalle  $[0, 1]$  de pas  $h = \frac{1}{N}$ , pour  $N$  égal à 5 et pour  $N$  égal à 15. Ces solutions approchées sont obtenues par la méthode d'Euler explicite, que l'on va définir plus bas, mais ce que nous cherchons à illustrer c'est ce que représente la solution approchée et non comment elle a été obtenue. On observe bien que pour la valeur de  $N$  la plus grande, il y a plus de points de discrétisation de l'intervalle  $[0, 1]$  (cette discrétisation est donc *plus précise*) et les valeurs approchées  $(y^0, \dots, y^N)$  sont *plus proches* des valeurs  $(y(t_0), \dots, y(t_N))$  de la solution exacte aux

points de la discrétisation. On s'attend bien à cela d'une méthode numérique : plus de points on utilise pour discrétiser l'intervalle  $[t_0, t_0 + T]$ , plus les valeurs approchées données par la méthode doivent être proches des valeurs exactes.



Une **méthode** ou **schéma numérique** est règle générale donnée par un algorithme qui, pour  $N \in \mathbb{N}$ , définit, pour  $n = 0, \dots, N - 1$ ,  $y^{n+1}$  en fonction de  $y^n$  ou en fonction de certaines des précédentes valeurs  $y^0, \dots, y^n$ , la valeur initiale  $y^0$  étant donnée. Ici on considérera toujours, comme mentionné plus haut,  $y^0 = y_0$  (mais on aurait pu considérer  $y^0 = y_N^0$ , avec  $y_N^0 \rightarrow y_0$ , lorsque  $N \rightarrow \infty$ ).

### 3.2.1 Exemples de schémas numériques.

Il y a essentiellement deux approches pour construire des schémas numériques. La première consiste d'abord à écrire l'équation  $y' = f(t, y)$  vérifiée par la solution exacte  $y$  de (PC) en un point  $t_n$  de la discrétisation. La fonction  $y : J \rightarrow \mathbb{R}^n$  vérifie *exactement*

$$y'(t_n) = f(t_n, y(t_n)).$$

On approche ensuite la dérivée  $y'(t_n)$  par un quotient aux différences faisant intervenir les valeurs de  $y$  en des points de la discrétisation. Par exemple, puisque

$$y'(t) \approx \frac{y(t+h) - y(t)}{h},$$

pour  $h$  petit, on a, comme  $t_{n+1} = t_n + h$ ,

$$y'(t_n) \approx \frac{y(t_{n+1}) - y(t_n)}{h}$$

et donc

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

Finalement on remplace les valeurs exactes (*mais inconnues, si on ne connaît pas la solution exacte*  $y(t_n)$ ) par des valeurs approchées  $y^n$ , pour  $n = 0, \dots, N$ , auxquelles on demande de vérifier l'équation ci-dessus non de manière approchée mais exacte. On cherche alors des valeurs approchées  $y^0, \dots, y^N$  vérifiant

$$\frac{y^{n+1} - y^n}{h} = f(t_n, y^n), \quad n = 0, \dots, N-1,$$

avec  $y^0 = y_0$  donné. Le schéma défini par cet algorithme s'appelle **schéma d'Euler explicite** (ou d'**Euler progressif** ou tout simplement d'Euler). Il peut être défini de manière équivalent par :

$$\begin{aligned} y^0 &= y_0, \\ y^{n+1} &= y^n + hf(t_n, y^n), \quad n = 0, \dots, N-1. \end{aligned}$$

**Remarque :** *il ne faut pas confondre les valeurs exactes  $y(t_n)$  de la solution exacte  $y$  aux instants  $t_n$  et les valeurs approchées  $y^n$ . Même si on sait que la solution exacte  $y$  existe, en général on ne sait pas la calculer et c'est à ça que les schémas numériques servent : à calculer des valeurs qui approchent des valeurs certes exactes mais inconnues.*

### Le schéma d'Euler implicite.

L'approximation

$$y'(t_{n+1}) \approx \frac{y(t_{n+1}) - y(t_n)}{h}$$

conduit, de la même manière, à partir de l'équation exacte

$$y'(t_{n+1}) = f(t_{n+1}, y(t_{n+1}))$$

vérifiée par la solution exacte  $y$ , au **schéma d'Euler implicite** ou **rétrograde**, défini par la relation

$$\begin{aligned} y^0 &= y_0, \\ y^{n+1} &= y^n + hf(t_{n+1}, y^{n+1}), \quad n = 0, \dots, N-1. \end{aligned}$$

On remarque que, pour calculer  $y^{n+1}$ , il faut résoudre par rapport à  $y^{n+1}$  l'équation à priori non linéaire  $y^{n+1} = y^n + hf(t_{n+1}, y^{n+1})$ . C'est pour cela que ce schéma s'appelle implicite :  $y^{n+1}$  est défini implicitement en fonction de  $y^n$ , alors que dans le schéma d'Euler explicite  $y^{n+1}$  est défini explicitement en fonction de  $y^n$ .

## Le schéma saute-mouton ou leap frog.

L'approximation centrée

$$y'(t_n) \approx \frac{y(t_{n+1}) - y(t_{n-1}))}{2h}$$

conduit, à partir de l'équation exacte

$$y'(t_n) = f(t_n, y(t_n))$$

vérifiée par la solution exacte  $y$ , au **schéma saute-mouton** ou **leapfrog**, défini par la relation

$$y^{n+1} = y^{n-1} + 2hf(t_n, y^n), \quad n = 1, \dots, N-1.$$

Dans ce schéma  $y^{n+1}$  est défini explicitement en fonction de  $y^n$  mais aussi de  $y^{n-1}$ . C'est ce qu'on appelle un schéma à deux pas. Sa mise en oeuvre (*son initialisation*) requiert de connaître  $y^0$  et  $y^1$ , ce dernier n'étant pas une donnée du problème. En général on définit  $y^1$  en utilisant un autre schéma numérique (par exemple le schéma d'Euler explicite, ou autre).

Une deuxième approche pour construire des schémas numériques consiste d'abord à intégrer l'équation  $y' = f(t, y)$ , vérifiée par la solution exacte  $y$  de (PC), dans un intervalle de la discrétisation  $[t_n, t_{n+1}[$ . On obtient

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

On approche ensuite l'intégrale

$$\int_{t_n}^{t_{n+1}} f(s, y(s)) ds$$

par une formule d'intégration numérique (*formule de quadrature*) faisant intervenir si possible les valeurs de  $y$  en des points de la discrétisation. Finalement on remplace les valeurs exactes  $y(t_n)$  par des valeurs approchées  $y^n$ . Par exemple, si l'on utilise la formule des rectangles à gauche pour approcher l'intégrale ci-dessus, on obtient

$$\int_{t_n}^{t_{n+1}} f(s, y(s)) ds \approx hf(t_n, y(t_n)).$$

En remplaçant les valeurs exactes  $y(t_n)$  par des valeurs approchées  $y^n$  dans l'approximation ci-dessus et dans la relation

$$y(t_{n+1}) - y(t_n) = \int_{t_n}^{t_{n+1}} f(s, y(s)) ds,$$

on obtient la relation

$$y^{n+1} = y^n + hf(t_n, y^n),$$

qui correspond à la méthode d'Euler explicite.

Si l'on utilise maintenant la formule des rectangles à droite pour approcher l'intégrale

$$\int_{t_n}^{t_{n+1}} f(s, y(s)) ds,$$

on obtient cette fois-ci la méthode d'Euler implicite.

### La méthode du point milieu.

Si l'on utilise la formule du point milieu pour approcher l'intégrale

$$\int_{t_n}^{t_{n+1}} f(s, y(s)) ds,$$

on obtient

$$y(t_{n+1}) - y(t_n) \approx hf\left(t_n + \frac{h}{2}, y\left(t_n + \frac{h}{2}\right)\right).$$

Comme le point  $t_n + \frac{h}{2}$  ne correspond pas à un point de la discrétisation, on approche aussi la valeur  $y\left(t_n + \frac{h}{2}\right)$  par une valeur approchée en suivant l'idée de la méthode d'Euler explicite :

$$y\left(t_n + \frac{h}{2}\right) \approx y(t_n) + \frac{h}{2}y'(t_n) = y(t_n) + \frac{h}{2}f(t_n, y(t_n)).$$

En remplaçant les valeurs exactes  $y(t_n)$  par des valeurs approchées  $y^n$  on obtient la méthode du point milieu

$$\begin{aligned} y^0 &= y_0, \\ y^{n+1} &= y^n + hf\left(t_n + \frac{h}{2}, y^n + \frac{h}{2}f(t_n, y^n)\right), \quad n = 0, \dots, N-1. \end{aligned}$$

### La méthode de Cranck-Nicholson.

En utilisant l'approximation

$$\int_{t_n}^{t_{n+1}} f(s, y(s)) ds \approx \frac{h}{2}(f(t_n, y(t_n)) + f(t_{n+1}, y(t_{n+1}))),$$

donnée par la formule des trapèzes, et en procédant comme avant, on obtient le schéma de Cranck-Nicholson

$$\begin{aligned} y^0 &= y_0, \\ y^{n+1} &= y^n + \frac{h}{2}(f(t_{n+1}, y^{n+1}) + f(t_n, y^n)), \quad n = 0, \dots, N-1. \end{aligned}$$

On appelle **méthode à un pas** une méthode numérique permettant de calculer  $y^{n+1}$  en fonction de la seule valeur antérieure  $y^n$ , pour  $n = 0, \dots, N-1$ . On appelle **méthode à  $r$  pas** une méthode numérique permettant de calculer  $y^{n+1}$  en fonction des  $r$  valeurs antérieures  $y^n, y^{n-1}, \dots, y^{n-r+1}$ . Une méthode à  $r$  pas requiert, pour sa mise en oeuvre, de définir les  $r$  premières valeurs  $y^0, \dots, y^{r-1}$  par éventuellement une autre méthode numérique.

On dit qu'une méthode est **explicite** lorsqu'elle permet de calculer  $y^{n+1}$  explicitement en fonction des valeurs antérieures. On dit qu'une méthode est **implicite** lorsque  $y^{n+1}$  est défini implicitement en fonction des valeurs antérieures.

On a vu dans cette section cinq exemples de schémas. Il y en a beaucoup d'autres, mais ces cinq exemples mettent déjà en évidence la multitude de situations que l'on peut avoir

lorsque l'on définit un schéma numérique : les schémas d'Euler explicite, saute-mouton et point milieu sont explicites, les schémas d'Euler implicite et de Cranck-Nicholson sont implicites. A l'exception du schéma saute-mouton, qui est un schéma à deux pas, tous les autres sont à un pas. Toutes ces différences influent dans la facilité de la mise en oeuvre d'un schéma, dans sa rapidité d'exécution (nombre d'opérations élémentaires que son exécution fait intervenir), dans sa précision par rapport à la solution exacte, et dans sa stabilité par rapport aux petites perturbations des données. On abordera ces deux dernières notions à la prochaine section.

**Remarque :** dans ces notes le pas de la discrétisation est noté  $h$  ; il est aussi très usuel de le noter  $\Delta t$ .

### 3.3 Théorie générale des méthodes explicites à un pas.

Nous allons nous concentrer sur une classe très large de schémas : les schémas explicites à un pas. Il existe de nombreuses méthodes numériques pour approcher une équation différentielle, et pas toutes sont dans cette classe. Mais un grand nombre de méthodes parmi les plus utilisées l'est, et on va donc étudier en détail la convergence de cette classe de méthodes.

Les méthodes explicites à un pas sont donc des méthodes pour lesquelles le calcul de  $y^{n+1}$  ne dépend que de  $y^n$  (et bien sûr aussi de  $h$  et de  $t^n$ , qui eux sont explicites). Ces méthodes s'écrivent sous la forme générique

$$(S) \quad y^{n+1} = y^n + h\Phi(t_n, y^n, h),$$

pour  $n = 0, \dots, N-1$ , avec  $y^0$  donné (nous considérerons toujours  $y^0 = y_0$ ), où  $\Phi : [t_0, t_0 + T] \times \mathbb{R}^n \times [0, 1] \longrightarrow \mathbb{R}^n$  est une fonction continue.

**Remarque :** Dans la formule (S) on a choisi l'intervalle  $[0, 1]$  pour la dépendance de  $\Phi$  par rapport à  $h$ . On aurait pu choisir un autre intervalle quelconque de la forme  $[0, h_{\max}]$ , avec  $h_{\max} > 0$ , puisque ce qui nous intéresse est le comportement de la suite  $(y^0, \dots, y^N)_N$  lorsque  $h \rightarrow 0$ .

Revenons sur les exemples vus à la section précédente. La **méthode d'Euler explicite** est bien une méthode qui s'écrit sous la forme (S) avec  $\Phi$  la fonction définie par

$$\Phi(t, y, h) = f(t, y).$$

Pour la **méthode du point milieu** on a

$$\Phi(t, y, h) = f\left(t + \frac{h}{2}, y + \frac{h}{2}f(t, y)\right).$$

Le schéma d'Euler implicite n'est pas *a priori* sous la forme (S). Mais on remarque que l'on a

$$\begin{aligned} y^{n+1} = y^n + hf(t_{n+1}, y^{n+1}) &\iff y^{n+1} - hf(t_{n+1}, y^{n+1}) = y^n \iff y^{n+1} - hf(t_n + h, y^{n+1}) = y^n \\ &\iff g(y^{n+1}, t_n, h) = y^n, \end{aligned}$$



où l'on a noté  $g(y, t, h)$  la fonction  $(y, t, h) \mapsto y - hf(t + h, y)$ .

Supposons que, pour  $(t, h) \in [t_0, t_0 + T] \times [0, 1]$  fixé, la fonction  $y \mapsto g(y, t, h)$  est inversible et notons dans ce cas par  $g^{-1}(y; t, h)$  sa réciproque. On a alors

$$y^{n+1} = y^n + hf(t_{n+1}, y^{n+1}) \iff y^{n+1} = g^{-1}(y^n; t_n, h) = y^n + h \frac{g^{-1}(y^n; t_n, h) - y^n}{h}.$$

Le schéma d'Euler implicite peut donc aussi s'écrire sous la forme (S), sous l'hypothèse que la fonction  $y \mapsto g(y, t, h)$  soit inversible. Si la fonction  $h \mapsto \frac{g^{-1}(y^n; t_n, h) - y^n}{h}$  admet un prolongement par continuité à  $h = 0$ , la fonction

$$\Phi(y, t, h) = \frac{g^{-1}(y; t, h) - y}{h}$$

est continue.

Notre objectif dans cette section est de trouver des conditions nécessaires sur la fonction  $\Phi$  pour que le schéma (S) *converge*, ce qui veut dire, dans un sens à préciser que  $\ll (y^0, \dots, y^N)$  tend vers  $(y(t_0), \dots, y(t_N)) \gg$  lorsque  $N \rightarrow +\infty$ , ou ce qui est équivalent, lorsque  $h \rightarrow 0$ . On va préciser le sens de cette limite et la notion de convergence d'un schéma au paragraphe suivant. *Il faut bien préciser le sens de cette limite car  $n$  n'est pas indépendant de  $N$  : on a  $t_n = t_0 + nh = t_0 + n \frac{T}{N}$ , les points  $t_n$  changent donc avec  $N$  et dire que  $y^n$  tend vers  $y(t_n)$  ne fait donc pas de sens.*

### 3.3.1 Convergence.

On commence par quelques définitions.

**Définition.**

**Erreur locale.**

Soit  $t \mapsto y(t)$  la solution exacte de (PC). On appelle **erreur locale à l'instant  $t_n$** , la quantité

$$e^n = y(t_n) - y^n.$$

**Remarque :** la quantité  $e^n$  dépend de  $h$  mais pour alléger les notations souvent on n'explicite pas cette dépendance.

**Définition.**

**Convergence.**

Le schéma défini par (S) est dit **convergent** sur l'intervalle  $[t_0, t_0 + T]$  si

$$\lim_{h \rightarrow 0} \left( \max_{n=0, \dots, N} \|e^n\| \right) = 0,$$

autrement dit si

$$\lim_{h \rightarrow 0} \left( \max_{n=0, \dots, N} \|y(t_n) - y^n\| \right) = 0.$$

Dans la définition ci-dessus,  $\|\cdot\|$  désigne une norme sur  $\mathbb{R}^n$ .

Il faut se rappeler que  $h$  et  $N$  sont liés par la relation  $h = \frac{T}{N}$ . Donc dire que  $h \rightarrow 0$  est équivalent à dire que  $N \rightarrow +\infty$ , et le schéma défini par  $(S)$  est convergent si et seulement si

$$\lim_{N \rightarrow +\infty} \left( \max_{n=0, \dots, N} \|y(t_n) - y^n\| \right) = 0.$$

Pour montrer la convergence d'un schéma numérique on doit alors :

- Calculer, pour  $N$  fixé, ou de manière équivalente pour  $h = \frac{T}{N}$  fixé, la solution approchée  $(y^0, \dots, y^N)$  associée à une discrétisation de l'intervalle  $[t_0, t_0 + T]$  de pas  $h = \frac{T}{N}$ .
- Calculer ensuite, pour  $N$  fixé, ou de manière équivalente pour  $h = \frac{T}{N}$  fixé, l'erreur globale

$$E_h = \max_{n=0, \dots, N} \|y(t_n) - y^n\|$$

(on remarque encore que  $N$  est lié à  $h$  ; plus  $h$  est petit, plus  $N$  est grand et plus il a de points  $t_n$  et de valeurs approchées  $y^n$ ). Pour chaque  $h$ ,  $E_h$  représente l'erreur maximale entre la solution exacte et la solution approchée associée à une discrétisation de l'intervalle  $[t_0, t_0 + T]$  de pas  $h$ .

On définit ainsi une suite  $(E_h)_{h>0}$  (c'est un abus de notation : il ne s'agit pas d'une application  $h > 0 \mapsto E_h$  mais de l'application  $N > 0 \mapsto h = \frac{T}{N} \mapsto E_h$  ; pour simplifier on note cette application par  $(E_h)_{h>0}$ , mais on garde à l'esprit que  $h$  prend des valeurs dans l'ensemble  $\{\frac{T}{N} \mid N > 0\}$ ).

- Montrer que la suite  $(E_h)_{h>0}$  converge vers 0, lorsque  $h \rightarrow 0$ .

On reprend ici l'exemple de l'équation  $y' = y$  pour illustrer les notions d'erreur locale et de convergence d'un schéma. On considère la fonction  $t \mapsto e^t$ , solution du problème de Cauchy  $y' = y$ ,  $y(0) = 1$ . Dans la première figure on considère la solution approchée obtenue avec le schéma d'Euler explicite pour une discrétisation avec  $N = 5$  points, ou, ce qui est équivalent, avec un pas  $h = 2/5$  (puisque l'on a fixé comme intervalle d'approximation l'intervalle  $[0, 2]$ , on a donc  $T = 2$  dans cet exemple).

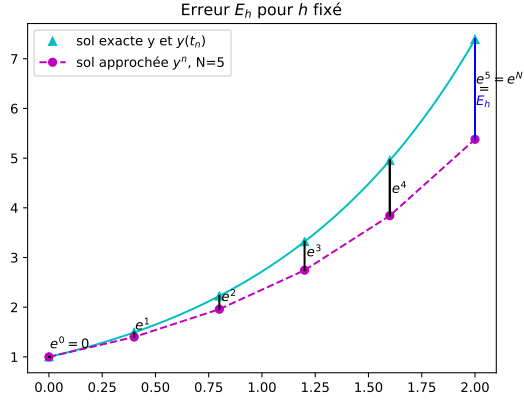


FIGURE 3.1 – Représentation de  $e^n$ ,  $n = 0, \dots, N = 5$ . On a dans cet exemple, pour la valeur de  $h = \frac{T}{N} = \frac{2}{5}$ ,  $E_h = \max\{|e^0|, \dots, |e^5|\} = |e^5|$ .

Dans la deuxième figure on représente la valeur de l'erreur globale  $E_h$  pour 3 valeurs de  $h$  différents. On observe bien que lorsque  $h$  diminue,  $E_h$  décroît (on s'attend à que  $E_h \rightarrow 0$ , lorsque  $h \rightarrow 0$ , mais il est aussi légitime de s'attendre à que  $E_h$  décroisse : si  $h$  diminue,  $N$  et donc le nombre de points de la discrétisation augmentent, on s'attend alors que l'approximation que l'on fait de  $y(t_n)$  soit plus précise et donc que la différence entre  $y^n$  et  $y(t_n)$  soit plus petite).

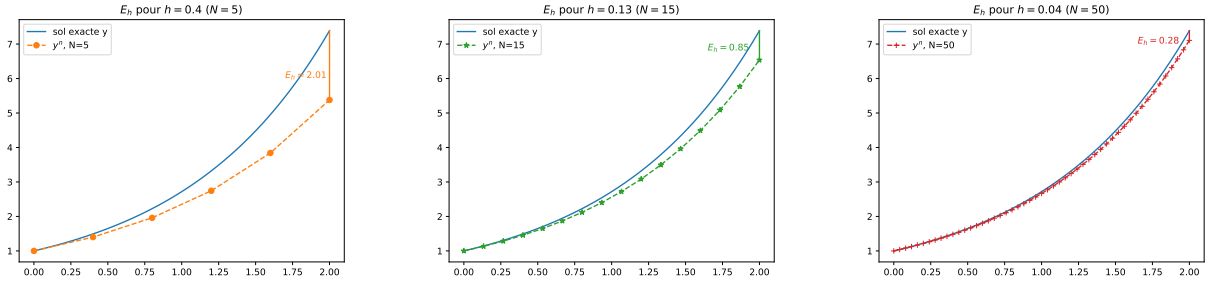


FIGURE 3.2 – Représentation de  $E_h$  pour les valeurs de  $h = 0.4$ ,  $0.13$  et  $0.04$  (correspondant respectivement aux valeurs de  $N = 5$ ,  $15$  et  $50$ ).

Dans la troisième figure, on représente les solutions approchées obtenues avec  $N = 5$ ,  $15$ ,  $50$  et  $1000$  points. On observe bien que le schéma d'Euler explicite semble converger, puisque la différence en valeur absolue entre la solution exacte et la solution approchée semble tendre vers 0.

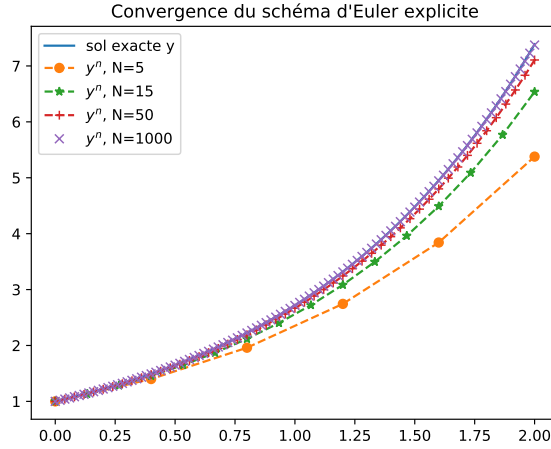


FIGURE 3.3 – Convergence du schéma d'Euler explicite : la solution approchée semble s'approcher de la solution exacte lorsque  $h$  diminue, ou de manière équivalente lorsque  $N$  augmente. La solution approchée obtenue avec  $N = 1000$  points semble se confondre avec la solution exacte. Si on pouvait augmenter la figure on verrait qu'il existe toujours une erreur entre les deux courbes, mais cette erreur est petite et elle deviendrait encore plus petite si on augmentait encore plus le nombre de points.

**Remarque :** Les notions d'erreur locale et de convergence que l'on a introduites sont aussi valables pour des schémas généraux, autres que les schémas explicites à un pas.

Les questions que l'on se pose par la suite sont les suivantes : quand est-ce qu'un schéma converge ? Qu'est-ce qu'empêche l'erreur de devenir grande ? Nous donnons dès maintenant un aperçu des réponses à ces questions :

1. Quand l'erreur commise par une itération du schéma est « petite », car la discrétisation utilisée est *cohérente* avec l'équation différentielle  $\rightarrow$  quand le schéma est *consistant* ;
2. Quand le schéma contrôle l'accumulation des erreurs, autrement dit quand le cumul des erreurs commises à chaque itération reste « petit »  $\rightarrow$  quand le schéma est *stable*.

La convergence d'un schéma est liée à ces deux notions, que nous allons étudier au paragraphe suivant : la **consistance** et la **stabilité**. Pour mieux comprendre dès maintenant ce qu'elles représentent, regardons comment évolue l'erreur dans le cas du schéma d'Euler explicite.

### Exemple - calcul de l'erreur pour le schéma d'Euler explicite.

On rappelle que le schéma d'Euler explicite est défini par la relation

$$y^{n+1} = y^n + hf(t_n, y^n), \quad n = 0, \dots, N-1,$$

avec  $y^0 = y_0$  donné. Cherchons à montrer que l'erreur locale à l'instant  $t_n$ ,  $e^n$ , vérifie également une relation de récurrence. On a  $e^{n+1} = y(t_{n+1}) - y^{n+1}$  et la suite des valeurs approchées

$(y^0, \dots, y^N)$  est définie par

$$y^{n+1} = y^n + hf(t_n, y^n), \quad n = 0, \dots, N-1, \quad y^0 = y_0.$$

Les valeurs exactes  $y(t_n)$ ,  $n = 0, \dots, N$ , ne vérifient pas la même relation mais une relation *approchée* :

$$\begin{aligned} y(t_{n+1}) &= y(t_n) + hf(t_n, y(t_n)) + \underbrace{y(t_{n+1}) - y(t_n) - hf(t_n, y(t_n))}_{:=\varepsilon^n}, \\ &= y(t_n) + hf(t_n, y(t_n)) + \varepsilon^n, \quad y(t_0) = y_0. \end{aligned}$$

Ici on a juste fait la somme et la différence de la même quantité, l'objectif étant de faire apparaître une récurrence similaire à celle vérifiée par les  $y^n$ .

La quantité  $\varepsilon^n$ , que l'on appellera *erreur de consistance* à l'étape  $n$ , on l'imagine *petite* car on a

$$\begin{aligned} y(t_{n+1}) &= y(t_n + h) \approx y(t_n) + hy'(t_n) \\ &= y(t_n) + hf(t_n, y(t_n)) \quad \text{car la fonction } t \mapsto y(t) \text{ est solution de } y' = f(t, y), \end{aligned}$$

et donc  $y(t_{n+1}) - y(t_n) - hf(t_n, y(t_n)) \approx 0$ , et d'autant plus petite que  $h$  est petit.

On a donc que la suite des valeurs exactes  $y(t_n)$ ,  $n = 0, \dots, N$ , dont la valeur initiale  $y_0$  est la même que pour la suite des valeurs approchées  $y^n$ ,  $n = 0, \dots, N$ , vérifie la même relation de récurrence que cette deuxième suite, entamée, à chaque itération, d'une erreur  $\varepsilon^n$ , censée être *petite*. On va voir qu'un schéma est consistant quand cette erreur est en effet petite. La stabilité d'un schéma est une propriété qui elle exige que la perturbation totale produite sur le schéma par les erreurs  $\varepsilon^n$  produites à chaque itération, reste de l'ordre du cumul de ces erreurs.

La suite des erreurs  $e^n$ ,  $n = 0, \dots, N$  vérifie donc :

$$\begin{aligned} e^{n+1} &= y(t_{n+1}) - y^{n+1} \\ &= y(t_n) + hf(t_n, y(t_n)) + \varepsilon^n - (y^n + hf(t_n, y^n)) \\ &= e^n + h(f(t_n, y(t_n)) - f(t_n, y^n)) + \varepsilon^n, \quad \forall n = 0, \dots, N-1. \end{aligned}$$

Nous faisons ici l'hypothèse que  $f$  est Lipschitzienne par rapport à la variable  $y$ , de constante de Lipschitz  $K$ . On a alors

$$\|f(t_n, y(t_n)) - f(t_n, y^n)\| \leq K\|y(t_n) - y^n\| = K\|e^n\|$$

et on obtient donc que

$$\|e^{n+1}\| \leq \|e^n\| + hK\|e^n\| + \|\varepsilon^n\| = \|e^n\|(1 + hK) + \|\varepsilon^n\|, \quad \forall n = 0, \dots, N-1.$$

On en déduit (on peut le montrer par récurrence sur  $n$ )

$$\begin{aligned}
\|e^{n+1}\| &\leq (1+hK)^{n+1} \underbrace{\|e^0\|}_{\|y_0-y^0\|=0} + (\|\varepsilon^n\| + (1+hK)\|\varepsilon^{n-1}\| + \dots + (1+hK)^n\|\varepsilon^0\|) \\
&= \|\varepsilon^n\| + (1+hK)\|\varepsilon^{n-1}\| + \dots + (1+hK)^n\|\varepsilon^0\| \\
&\leq (\|\varepsilon^n\| + \|\varepsilon^{n-1}\| + \dots + \|\varepsilon^0\|) \max\{1, (1+hK), \dots, (1+hK)^n\} \\
&= (\|\varepsilon^n\| + \|\varepsilon^{n-1}\| + \dots + \|\varepsilon^0\|)(1+hK)^n \\
&\leq (\|\varepsilon^n\| + \|\varepsilon^{n-1}\| + \dots + \|\varepsilon^0\|)(1+hK)^N, \quad \text{car } n \leq N, \\
&\leq (\|\varepsilon^n\| + \|\varepsilon^{n-1}\| + \dots + \|\varepsilon^0\|) \underbrace{(e^{hK})^N}_{e^{\frac{T}{N}KN} = e^{KT}}, \quad \text{car } e^{hK} = 1 + hK + \frac{(hK)^2}{2} + \dots, \\
&\leq e^{KT} (\|\varepsilon^n\| + \|\varepsilon^{n-1}\| + \dots + \|\varepsilon^0\|), \quad \forall n = 0, \dots, N-1.
\end{aligned}$$

L'inégalité que l'on vient d'obtenir illustre la propriété de stabilité du schéma d'Euler explicite : sous l'hypothèse que la fonction  $f$  soit Lipschitzienne, l'erreur entre la solution approchée et la solution exacte est contrôlée par la somme des erreurs commises à chaque itération.

Comme  $T$  est fixé, si  $(\|\varepsilon^n\| + \|\varepsilon^{n-1}\| + \dots + \|\varepsilon^0\|) \rightarrow 0$ , lorsque  $h \rightarrow 0$ , alors le schéma d'Euler explicite est convergent.

La quantité  $\varepsilon^n$  est appelée erreur de consistance à l'étape  $n$ . On dira qu'un schéma est consistant quand la somme ci-dessus tend vers 0, lorsque  $h \rightarrow 0$ . Nous montrerons à la section suivante que c'est bien le cas pour le schéma d'Euler explicite, sous l'hypothèse que la solution  $y$  de (PC) soit de classe  $C^2$ .

Nous allons maintenant préciser la définition des notions de consistance et de stabilité.

### 3.3.2 Consistance des schéma à un pas.

Le schéma (S) définit la suite des solutions approchées  $(y^0, \dots, y^N)$  par la relation

$$y^{n+1} = y^n + h\Phi(t_n, y^n, h), \quad n = 0, \dots, N-1,$$

avec  $y^0 = y_0$  donné. La suite  $(y(t_0), \dots, y(t_N))$  des valeurs de la solution exacte aux instants  $t_n$  ne vérifie pas la même relation mais on espère qu'elle la vérifie de manière approchée (car le schéma numérique est censé traduire de manière approchée ce qui se passe avec la solution exacte  $y$ ). On espère alors que la quantité  $y(t_{n+1}) - y(t_n) - h\Phi(t_n, y(t_n), h)$  soit petite, en norme.

**Définition.**

**Erreur de consistance locale.**

Soit  $y$  une solution de l'équation différentielle  $y' = f(t, y)$ .

On appelle **erreur de consistance** du schéma  $(S)$  à l'étape  $n \in \{0, \dots, N-1\}$ , relative à la solution  $y$ , la quantité  $\varepsilon^n$  définie par

$$\varepsilon^n = y(t_{n+1}) - y(t_n) - h\Phi(t_n, y(t_n), h).$$

L'erreur de consistance est l'erreur produite par une étape du schéma lorsque l'on applique le schéma à la solution exacte  $y(t_n)$  à l'instant  $t_n$  (*c'est bien la différence entre la solution exacte  $y(t_{n+1})$  à l'instant  $t_{n+1}$  et le résultat d'une étape du schéma appliqué à  $y(t_n)$* ).

**Remarque :** l'erreur de consistance dépend de  $h$  et de la solution  $y$  de l'équation  $y' = f(t, y)$  que l'on considère, même si cette dépendance n'est pas prise en compte dans la notation utilisée. On ne l'écrira pas mais on garde à l'esprit que  $\varepsilon^n = \varepsilon^n(h, y)$ .

### Définition.

#### Schéma consistant. Ordre de consistance.

Le schéma  $(S)$  est dit **consistant** avec l'équation différentielle  $y' = f(t, y)$  si pour toute solution  $y$  de cette équation, on a

$$\lim_{h \rightarrow 0} \sum_{n=0}^{N-1} \|\varepsilon^n\| = 0.$$

Soit  $p \in \mathbb{N}^*$ . Le schéma  $(S)$  est dit **consistant d'ordre au moins  $p$**  avec l'équation différentielle  $y' = f(t, y)$  si pour toute solution  $y$  de cette équation, il existe une constante  $C > 0$ , indépendante de  $h$ , tel que

$$\forall N \in \mathbb{N}, \quad \sum_{n=0}^{N-1} \|\varepsilon^n\| \leq Ch^p.$$

Le schéma est dit **consistant d'ordre  $p$**  s'il est consistant d'ordre  $p$  mais n'est pas consistant d'ordre au moins  $p+1$ .

**Remarque :** la constante  $C$  ne dépend pas de  $h$  mais elle dépendra en général fortement de la solution  $y$  de l'équation  $y' = f(t, y)$ .

**Remarque :** On utilise souvent la formulation « schéma d'ordre  $p$  » au lieu de « schéma consistant d'ordre  $p$  ». Un schéma d'ordre  $p$  est donc un schéma dont l'erreur de consistance vérifie la propriété ci-dessus. On verra un peu plus loin que cela implique que l'erreur globale du schéma vérifie la même estimation.

Pour montrer la consistance à l'ordre au moins  $p$  d'un schéma (la consistance d'un schéma étant conséquence de la consistance à l'ordre au moins 1), il suffit de montrer une estimation sur l'erreur de consistance locale de la forme

$$\|\varepsilon^n\| \leq Ch^{p+1},$$

avec  $C$  une constante indépendante de  $h$ . En effet, si on a une telle estimation, on obtient

$$\sum_{n=0}^{N-1} \|\varepsilon^n\| \leq \sum_{n=0}^{N-1} Ch^{p+1} = Ch^{p+1}N = Ch^{p+1}\frac{T}{h} = \tilde{C}h^p,$$

(on a utilisé que  $h = \frac{T}{N}$  donc  $N = \frac{T}{h}$ ), avec  $\tilde{C} = CT$ . La constante  $\tilde{C}$  est bien indépendante de  $h$  car le temps final  $t_0 + T$  est fixé.

### Exemple (suite) - consistance du schéma d'Euler explicite.

Montrons que si  $y$  est une solution de l'équation  $y' = f(t, y)$  de classe  $C^2$ , alors l'erreur de consistance locale du schéma d'Euler explicite, relative à  $y$ , est en  $\mathcal{O}(h^2)$ . Plus précisément, montrons qu'il existe une constante  $C > 0$ , indépendante de  $h$ , tel que

$$\forall N \in \mathbb{N}, \quad \|\varepsilon^n\| \leq Ch^2, \quad \forall n = 0, \dots, N-1.$$

On a

$$\begin{aligned} \varepsilon^n &= y(t_{n+1}) - y(t_n) - hf(t_n, y(t_n)) \\ &= y(t_n + h) - y(t_n) - hf(t_n, y(t_n)). \end{aligned}$$

Comme  $y$  est de classe  $C^2$ , on peut développer à l'ordre 2  $y(t_n + h)$  autour de  $t_n$  et on obtient qu'il existe  $\xi_n$  entre  $t_n$  et  $t_n + h$  tel que

$$\begin{aligned} \varepsilon^n &= y(t_n) + hy'(t_n) + \frac{h^2}{2}y''(\xi_n) - y(t_n) - hf(t_n, y(t_n)) \\ &= \cancel{y(t_n)} + \cancel{hy'(t_n)} + \frac{h^2}{2}y''(\xi_n) - \cancel{y(t_n)} - \cancel{hy'(t_n)} \quad \text{car } y \text{ vérifie } y'(t_n) = f(t_n, y(t_n)) \\ &= \frac{h^2}{2}y''(\xi_n). \end{aligned}$$

On a donc

$$\|\varepsilon^n\| \leq Ch^2,$$

avec

$$C = \frac{1}{2} \max_{t \in [t_0, t_0+T]} \|y''(t)\|$$

une constante positive indépendante de  $h$ .



Supposons alors  $f$  une fonction de classe  $C^1$ . Dans ce cas toute solution  $y$  de l'équation  $y' = f(t, y)$  est de classe  $C^2$  et on conclut que le schéma d'Euler explicite est consistant à l'ordre au moins 1 avec l'équation différentielle  $y' = f(t, y)$ , car on a

$$\forall N, \sum_{n=0}^{N-1} \|\varepsilon^n\| \leq N \frac{h^2}{2} \max_{t \in [t_0, t_0+T]} \|y''(t)\| = \left( \frac{T}{2} \max_{t \in [t_0, t_0+T]} \|y''(t)\| \right) h$$

(on a à nouveau utilisé  $N = \frac{T}{h}$  ; la constante  $C$  que l'on cherche doit être indépendante de  $h$  donc aussi de  $N$ ).

Pour finir ce paragraphe nous donnons un critère très utile qui permet de montrer la consistance d'un schéma défini par la relation (S). Nous donnons la preuve de ce résultat à titre indicatif.

**Proposition.**

**Condition suffisante de consistance.**

Soit l'équation différentielle  $y' = f(t, y)$  et un schéma numérique de la forme (S) pour approcher le problème de Cauchy (PC) pour cette équation.

Supposons  $f$  une fonction continue et localement Lipschitzienne par rapport à sa deuxième variable et  $\Phi$  une fonction continue.

Alors, si

$$\Phi(t, y, 0) = f(t, y), \quad \text{pour tout } (t, y) \in I \times \mathbb{R}^n,$$

le schéma défini par (S) est consistant avec l'équation  $y' = f(t, y)$ .

*Démonstration.* Soit  $y$  une solution de  $y' = f(t, y)$ . On veut montrer que

$$\lim_{h \rightarrow 0} \sum_{n=0}^{N-1} \|\varepsilon^n\| = 0,$$

i. e. que

$$\forall \delta > 0, \exists \epsilon > 0 : 0 < h < \epsilon \implies \sum_{n=0}^{N-1} \|\varepsilon^n\| < \delta.$$

On a

$$\begin{aligned} \varepsilon^n &= y(t_{n+1}) - y(t_n) - h\Phi(t_n, y(t_n), h) \\ &= \int_{t_n}^{t_{n+1}} \left( y'(s) - \Phi(t_n, y(t_n), h) \right) ds = \int_{t_n}^{t_{n+1}} \left( f(s, y(s)) - \Phi(t_n, y(t_n), h) \right) ds \\ &= \int_{t_n}^{t_{n+1}} \left( f(s, y(s)) - f(t_n, y(t_n)) \right) + \left( \underbrace{\Phi(t_n, y(t_n), 0)}_{f(t_n, y(t_n)), \text{ par hypothèse}} - \Phi(t_n, y(t_n), h) \right) ds. \end{aligned}$$

Soit  $\delta > 0$ .

Comme la fonction  $h \mapsto \Phi(t_n, y(t_n), h)$  est continue en  $h = 0$ , on a

$$\exists \epsilon_1 > 0 : 0 < h < \epsilon_1 \implies \|\Phi(t_n, y(t_n), 0) - \Phi(t_n, y(t_n), h)\| < \frac{\delta}{2T}.$$

Comme la fonction  $s \mapsto f(s, y(s))$  est continue sur  $[t_0, t_0 + T]$ , elle est uniformément continue sur cet intervalle et

$$\exists \epsilon_2 > 0 : s, t \in [t_0, t_0 + T], |s - t| < \epsilon_2 \implies \|f(s, y(s)) - f(t, y(t))\| < \frac{\delta}{2T}.$$

On a en particulier que, si  $h < \epsilon_2$ , alors pour tout  $s \in [t_n, t_{n+1}]$ ,

$$\|f(s, y(s)) - f(t_n, y(t_n))\| < \frac{\delta}{2T}.$$

Soit alors  $\epsilon = \min\{\epsilon_1, \epsilon_2\}$ . Si  $0 < h < \epsilon$ ,

$$\begin{aligned} \|\varepsilon^n\| &= \left\| \int_{t_n}^{t_{n+1}} \left( f(s, y(s)) - f(t_n, y(t_n)) \right) + \left( \Phi(t_n, y(t_n), 0) - \Phi(t_n, y(t_n), h) \right) ds \right\| \\ &\leq \int_{t_n}^{t_{n+1}} \|f(s, y(s)) - f(t_n, y(t_n))\| + \|\Phi(t_n, y(t_n), 0) - \Phi(t_n, y(t_n), h)\| \\ &\leq h \frac{\delta}{2T} + h \frac{\delta}{2T} = h \frac{\delta}{T}. \end{aligned}$$

On en déduit que

$$\sum_{n=0}^{N-1} \|\varepsilon^n\| \leq N h \frac{\delta}{T} = \delta.$$

□

**Remarque :** le théorème précédent donne une condition suffisante de consistance d'un schéma. Il ne nous dit par contre rien sur l'ordre de consistance du schéma.

On aurait pu utiliser le théorème précédent pour montrer la consistance du schéma d'Euler explicite. Pour le schéma d'Euler explicite on a  $\Phi(t, y, h) = f(t, y)$  pour tout  $h$ , en particulier pour  $h = 0$ . Mais les calculs que l'on a fait sur l'erreur de consistance locale pour le schéma d'Euler explicite nous donnent plus d'information que la simple consistance du schéma : il nous disent que le schéma d'Euler explicite est consistant d'ordre au moins égal à 1.

Nous parlons de l'intérêt de connaître l'ordre de consistance d'un schéma. Cet ordre de consistance quantifie en quelque sort la qualité de l'approximation effectuée. Aussi, nous verrons qu'un schéma qui est consistant à l'ordre au moins  $p$ , converge aussi à l'ordre au moins  $p$ . Cela veut dire que l'erreur  $E_h$  va tendre vers 0, quand  $h \rightarrow 0$ , au moins comme  $h^p$ . Ceci est important car plus  $p$  est grand, plus la convergence est rapide.

### 3.3.3 Stabilité des schéma à un pas.

**Définition.****Schéma stable.**

Le schéma numérique défini par  $(S)$  est dit **stable** s'il existe une constante  $M > 0$ , indépendante de  $h$ , tel que pour toutes suites  $y^n$  et  $\tilde{y}^n$  définies par

$$\begin{aligned} y^{n+1} &= y^n + h\Phi(t_n, y^n, h), \quad n = 0, \dots, N-1, \quad y^0 \text{ donné}, \\ \tilde{y}^{n+1} &= \tilde{y}^n + h\Phi(t_n, \tilde{y}^n, h) + \rho^n, \quad n = 0, \dots, N-1, \quad \tilde{y}^0 \text{ donné}, \end{aligned}$$

avec  $\rho^n \in \mathbb{R}$ ,  $n = 0, \dots, N-1$ , on ait

$$\max_{n=0, \dots, N} \|y^n - \tilde{y}^n\| \leq M \left( \|y^0 - \tilde{y}^0\| + \sum_{n=0}^{N-1} \|\rho^n\| \right).$$

La constante  $M$  est appelée constante de stabilité du schéma.

Un schéma stable est alors un schéma tel que, lorsque l'on introduit une perturbation  $\rho^n$  à chaque étape du schéma, l'erreur totale du schéma est contrôlée par le cumul des perturbations.

**Exemple - suite : stabilité du schéma d'Euler explicite**

Nous l'avons déjà montré, lorsque nous avons calculé l'erreur pour le schéma d'Euler explicite : sous l'hypothèse que  $f$  soit Lipschitzienne par rapport à sa deuxième variable, de constante de Lipschitz  $K$ , le schéma d'Euler explicite est stable. En effet, si  $y^n$  et  $\tilde{y}^n$  sont deux suites vérifiant

$$\begin{aligned} y^{n+1} &= y^n + hf(t_n, y^n), \quad n = 0, \dots, N-1, \quad y^0 \text{ donné}, \\ \tilde{y}^{n+1} &= \tilde{y}^n + hf(t_n, \tilde{y}^n) + \rho^n, \quad n = 0, \dots, N-1, \quad \tilde{y}^0 \text{ donné}, \end{aligned}$$

on obtient, en reproduisant les mêmes calculs que pour l'erreur  $e^n$ ,

$$\|y^{n+1} - \tilde{y}^{n+1}\| \leq \|y^n - \tilde{y}^n\|(1 + hk) + \|\rho^n\| \leq e^{KT} \left( \|y^0 - \tilde{y}^0\| + \sum_{p=0}^n \|\rho^p\| \right).$$

On a donc

$$\max_{n=0, \dots, N} \|y^n - \tilde{y}^n\| \leq M \left( \|y^0 - \tilde{y}^0\| + \sum_{n=0}^{N-1} \|\rho^n\| \right),$$

avec  $M = e^{KT}$  constante positive indépendante de  $h$ .

La proposition suivante donne une condition suffisante de stabilité d'un schéma numérique à un pas.

**Proposition.****Condition suffisante de stabilité.**

Soit un schéma numérique de la forme (S).

S'il existe une constante  $K > 0$  tel que pour tous  $y_1, y_2 \in \mathbb{R}^n$ ,  $t \in [t_0, t_0 + T]$  et  $h \in [0, 1]$  on ait

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

alors le schéma est stable.

**Remarque :** la condition  $\|\Phi(t, y_1, h) - \Phi(t, y_2, h)\| \leq K\|y_1 - y_2\|$ , pour tous  $y_1, y_2 \in \mathbb{R}^n$ , pour tous  $t \in [t_0, t_0 + T]$ ,  $h \in [0, 1]$ , correspond à demander que  $\Phi$  soit Lipschitzienne par rapport à la variable  $y$  dans  $\mathbb{R}^n$ .

*Démonstration.* Soient  $(\rho^n)_{n=0, \dots, N}$  une suite et  $y^n, \tilde{y}^n$  définies par

$$\begin{aligned} y^{n+1} &= y^n + h\Phi(t_n, y^n, h), \quad n = 0, \dots, N-1, \\ \tilde{y}^{n+1} &= \tilde{y}^n + h\Phi(t_n, \tilde{y}^n, h) + \rho^n, \quad n = 0, \dots, N-1. \end{aligned}$$

On a alors

$$\begin{aligned} \|y^{n+1} - \tilde{y}^{n+1}\| &\leq \|y^n - \tilde{y}^n\| + h\|\Phi(t_n, y^n, h) - \Phi(t_n, \tilde{y}^n, h)\| + \|\rho^n\| \\ &\leq \|y^n - \tilde{y}^n\| + hK\|y^n - \tilde{y}^n\| + \|\rho^n\| \\ &= \|y^n - \tilde{y}^n\|(1 + hK) + \|\rho^n\|. \end{aligned}$$

On a donc (revoir les calculs faits pour l'erreur  $e^n$  du schéma d'Euler explicite)

$$\|y^{n+1} - \tilde{y}^{n+1}\| \leq e^{KT} \left( \|y^0 - \tilde{y}^0\| + \sum_{p=0}^n \|\rho^p\| \right)$$

et donc

$$\max_{n=0, \dots, N} \|y^n - \tilde{y}^n\| \leq e^{KT} \left( \|y^0 - \tilde{y}^0\| + \sum_{p=0}^{N-1} \|\rho^p\| \right).$$

□

**Remarque :** En pratique, la fonction  $\Phi$  est rarement globalement Lipschitzienne. Mais elle est souvent de classe  $C^1$  par rapport à  $y$ , si  $f$  l'est aussi. Dans ce cas, le résultat de la proposition reste vrai si l'on peut montrer que la solution approchée donnée par le schéma prend ses valeurs dans un compact (puisque dans ce compact  $\Phi$  va alors être globalement Lipschitzienne). L'image de l'intervalle  $[t_0, t_0 + T]$  par la solution exacte, qui est une fonction continue, est bien un compact. Mais rien ne nous dit que la suite des solutions approchées donnée par le schéma prend aussi ses valeurs dans un compact, cela doit être analysé au cas par cas.

**Exemple - suite : convergence du schéma d'Euler explicite.**

Supposons  $f$  une fonction de classe  $C^1$  et globalement Lipschitzienne par rapport à sa deuxième variable  $y$ . Nous avons montré que comme  $f$  est globalement Lipschitzienne en  $y$ , le schéma d'Euler explicite est stable et donc la suite des erreurs  $e^n$ ,  $n = 0, \dots, N$  vérifie

$$\|e^{n+1}\| \leq e^{KT} \sum_{p=0}^n \|\varepsilon^p\|, \quad \forall n = 0, \dots, N-1,$$

où  $\varepsilon^n$  est l'erreur de consistance à l'étape  $n$  du schéma. On a alors

$$\max_{n=0, \dots, N} \|e^n\| \leq e^{KT} \sum_{n=0}^{N-1} \|\varepsilon^n\|.$$

Comme le schéma est consistant avec l'équation  $y' = f(t, y)$ , on a

$$\sum_{n=0}^{N-1} \|\varepsilon^n\| \xrightarrow{h \rightarrow 0} 0.$$

On conclut alors que

$$\lim_{h \rightarrow 0} \max_{n=0, \dots, N} \|e^n\| = 0,$$

autrement dit le schéma d'Euler explicite est convergent.

### 3.3.4 Le théorème de Lax.

La démarche que l'on a utilisé pour montrer la convergence du schéma d'Euler explicite s'étend à n'importe quel autre schéma explicite à un pas de la forme (S), comme le montre la preuve du théorème suivant, connu sous le nom de théorème de Lax.

**Theoreme. (Lax).**

Si le schéma défini par (S) est stable et consistant, alors il est convergent.

*Démonstration.* La solution approchée  $(y^0, \dots, y^N)_N$  vérifie

$$y^{n+1} = y^n + h\Phi(t_n, y^n, h), \quad n = 0, \dots, N-1, \quad y^0 = y_0,$$

alors que la suite  $(y(t_0), \dots, y(t_N))$  des valeurs de la solution exacte aux instants  $t_n$  vérifie

$$\begin{aligned} y(t_{n+1}) &= y(t_n) + h\Phi(t_n, y(t_n), h) + y(t_{n+1}) - y(t_n) - h\Phi(t_n, y(t_n), h) \\ &= y(t_n) + h\Phi(t_n, y(t_n), h) + \varepsilon^n, \quad n = 0, \dots, N-1, \quad y(t_0) = y_0, \end{aligned}$$

avec  $\varepsilon^n$  l'erreur de consistance du schéma, relative à la solution exacte  $y$ , à l'étape  $n$ . Comme le schéma est stable, on obtient

$$\max_{n=0,\dots,N} \|y(t_n) - y^n\| \leq M \left( \underbrace{\|y(t_0) - y^0\|}_0 + \sum_{n=0}^{N-1} \|\varepsilon^n\| \right) = M \sum_{n=0}^{N-1} \|\varepsilon^n\|,$$

avec  $M$  la constante de stabilité du schéma. Comme le schéma est consistant et  $M$  est indépendante de  $h$ , on a

$$\lim_{h \rightarrow 0} M \sum_{n=0}^{N-1} \|\varepsilon^n\| = 0,$$

donc

$$\lim_{h \rightarrow 0} \max_{n=0,\dots,N} \|y(t_n) - y^n\| = 0.$$

□

### 3.3.5 Ordre d'une méthode, estimation de l'erreur.

La consistance et la stabilité d'une méthode permettent de montrer sa convergence, mais ne donnent pas d'information sur comment l'erreur varie avec le pas  $h$ . La notion d'ordre d'une méthode permet de quantifier comment l'erreur se comporte en fonction de  $h$  et de donner une information sur la vitesse à laquelle le schéma converge.

L'intérêt de la notion d'ordre (de consistance) est qu'elle permet d'obtenir une estimation du même ordre  $p$  sur l'erreur globale de la méthode. C'est ce qu'établit le résultat suivant, qui est un raffinement du théorème de Lax.

#### **Proposition. Estimation d'erreur.**

Si le schéma défini par  $(S)$  est stable et consistant à l'ordre au moins  $p$ , alors il est convergent. En plus, il existe une constante  $C > 0$ , indépendante de  $h$ , tel que

$$\max_{n=0,\dots,N} \|y(t_n) - y^n\| \leq Ch^p.$$

Pour montrer ce théorème, il faut reprendre la preuve du théorème de Lax en remplaçant l'erreur de consistance globale par son estimation d'ordre  $p$ .

**Remarques :**

1. *L'intérêt d'une méthode d'ordre  $p$  par rapport à une méthode d'ordre  $q$  inférieur à  $p$  est que, asymptotiquement, sa précision est bien meilleure, puisque  $h^p$  converge vers 0 plus rapidement que  $h^q$ . On s'attend que, à pas de discrétisation  $h$  fixé, l'erreur associée à un schéma d'ordre  $p$  soit  $1/h^{p-q}$  fois plus petite que l'erreur associée à un schéma d'ordre  $q$ . Il faut cependant faire attention que l'estimation d'erreur du théorème précédent fait intervenir une constante  $C$  inconnue, qui dépend de la norme des dérivées de la solution  $y$  et qui peut être grande. Cette constante dépend du schéma, elle n'est pas la même pour un schéma d'ordre  $p$  et pour un schéma d'ordre  $q$ . La précision supérieure à pas égal pour un schéma d'ordre supérieur n'est donc pas garantie, mais elle est le plus souvent observée.*
2. *Supposons la constante  $C$  de l'ordre de l'unité pour un schéma d'ordre  $p = 1$ . Pour obtenir une approximation avec une précision de  $10^{-4}$ , il faut prendre un pas  $h = 10^{-4}$ . Si cette constante  $C$  reste de l'ordre de l'unité pour un schéma d'ordre  $p + 1 = 2$ , pour obtenir la même précision, il suffit de prendre un pas  $h$  tel que  $h^2 = 10^{-4}$ , soit  $h = 10^{-2}$ , autrement dit un pas  $h$  100 fois plus grand. Cela veut dire que par rapport à un schéma d'ordre 1, pour un schéma d'ordre 2 on obtient la même précision en utilisant une discrétisation avec 100 fois moins de points, il y a donc 100 fois moins de valeurs  $y^0, y^1, \dots, y^N$  à calculer. En revanche, pour une méthode d'ordre supérieur, le calcul de chaque  $y^n$  est règle générale plus coûteux. Il faut prendre en compte que plus l'ordre d'une méthode est élevée, plus le coût de cette méthode (nombre d'opérations pour son exécution, temps de calcul, place mémoire occupée) est grand; il faut alors au cas par cas trouver un équilibre entre le gain de précision et le coût de la méthode, lorsque l'on doit décider de l'utilisation d'une méthode d'ordre plus élevé.*
3. *Une des plus connues et utilisées méthodes d'ordre élevé est la méthode de Runge-Kutta 4 (RK4), vue au TP 3. Il s'agit, comme son nom l'indique, d'une méthode d'ordre 4. Les méthodes dites de Runge-Kutta sont des méthodes qui se basent aussi sur l'intégration de l'EDO entre les instants  $t_n$  et  $t_{n+1}$ , mais qui utilisent des instants intermédiaires entre  $t_n$  et  $t_{n+1}$  pour le calcul de la valeur approchée en  $t_{n+1}$ . Ce sont des schémas à un pas (même si l'écriture d'un pas du schéma RK4 peut paraître bien plus compliquée que celle d'autres méthodes à un pas que l'on a vu).*