

→ Schémas numériques.

Une solution exacte c'est toujours sympa,
mais pour nombreux, une approximation suffira.

■ La motivation.

D – Comment tu ferais pour résoudre $x''(t) = x^2(t) + 6t$?

T – C'est une équation de Painlevé. On ne peut pas en donner une expression avec des fonctions élémentaires. Ne t'inquiètes pas, c'est le cas de beaucoup d'équations différentielles.

D – Oui mais si on a besoin de la valeur en un certain instant ?

T – On approxime la solution. C'est comme ça que les ordinateurs tracent les graphes de ces fonctions.

■ Rappels.

Une fonction $f : E \rightarrow F$ est dite **K -Lipschitz** (ou lipschitzienne de constante de Lipschitz K) si

$$\forall x, y \in E, d(f(x), f(y)) \leq K \cdot d(x, y)$$

La plupart du temps, on écrit $d(x, y)$ pour $\|x - y\|$. Une fonction est **strictement contractante** si elle est K -Lipschitz avec $K < 1$.

Par exemple, soit le problème

$$(x'(t), y'(t)) = \left(\frac{x(t) + y(t)}{2}, \frac{x(t) - y(t)}{2} \right)$$

La fonction $f(t, (x, y))$ utilisée ici est $f(t, (x, y)) = \left(\frac{x+y}{2}, \frac{x-y}{2} \right)$. On vérifie que

$$\left\| \left(\frac{x+y}{2}, \frac{x-y}{2} \right) - \left(\frac{u+v}{2}, \frac{u-v}{2} \right) \right\| \leq \frac{1}{\sqrt{2}} \|(x, y) - (u, v)\|$$

Donc f est $\frac{1}{\sqrt{2}}$ -lipschitzienne en espace (et donc contractante, mais il n'y en a pas besoin ici). Cela garantit l'existence d'une unique solution au problème.

■ Problème de Cauchy.

On ramène tous les problèmes à cette forme, qu'on appelle un **problème de Cauchy** :

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(0) = x_0 \end{cases}$$

avec $f : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ continue sur $I \times \mathbb{R}^m$ avec I intervalle de \mathbb{R} contenant 0. On veut trouver une solution $x : I \rightarrow \mathbb{R}^m$ qui est \mathcal{C}^1 . Si jamais f est Lipschitz en espace, le théorème de Cauchy-Lipschitz global garantit l'existence et l'unicité d'une solution sur I . Le cas échéant, on peut utiliser le théorème de Cauchy-Lipschitz local.

■ Discretisation.

On doit, pour appliquer une méthode/un schéma numérique, diviser l'intervalle de temps étudié $([0, T])$ en une liste de temps ; on discrétise. Typiquement, on prendra la suite $t_n = t_0 + n \cdot h$, avec $h > 0$ le pas de temps, et $t_0 \in [0, T]$ un point de départ. On va jusqu'à $N_h = \lfloor \frac{T-t_0}{h} \rfloor$ (donc le plus grand entier qui donne $t_{N_h} \leq T$). Au pire, pour une subdivision plus compliquée, on peut toujours récupérer les pas de temps $h_n := t_{n+1} - t_n$.

■ Cauchy-Lipschitz local.

Soit \mathcal{U} ouvert de $\mathbb{R} \times \mathbb{R}^m$, et f continue de \mathcal{U} vers \mathbb{R}^m , localement lipschitzienne en espace sur \mathcal{U} . Alors pour tout $(t_0, x_0) \in \mathcal{U}$, il existe I intervalle contenant t_0 et un unique $x \in \mathcal{C}^1(I, \mathbb{R}^m)$ tel que pour $t \in I$, $(t, x(t)) \in \mathcal{U}$ et

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

Rappelons que f "localement lipschitzienne en espace sur \mathcal{U} " signifie que $\forall (\tilde{t}, \tilde{x}) \in \mathcal{U}$, il existe un voisinage $\mathcal{V} \subset \mathcal{U}$ qui contient (\tilde{t}, \tilde{x}) et une constante $C > 0$ tels que $\|f(t, y) - f(t', z)\| \leq C \|y - z\|$ pour tous (t, y) et (t', z) dans \mathcal{V} . Autrement dit, partout dans \mathcal{U} il y a un voisinage sur lequel f est Lipschitz.

Pour plus de détails sur ce théorème et ces conditions, voir les fiches (ou le cours) de Calcul Différentiel et Équations Différentielles.

■ Méthode numérique.

On peut enfin définir une méthode/un schéma numérique : c'est une procédure qui donne, en fonction de f , de $(t_n)_{n \in \{0, \dots, N\}}$ et de x_0 , une suite de valeurs $(x_n)_{n \in \{0, \dots, N\}}$ qui approxime $x(t_n)$. On peut opérer deux distinctions fondamentales.

Une méthode est dite **à un pas** si on peut l'écrire comme $x_{n+1} = x_n + h \cdot \phi(t_n, x_n, h)$; c'est-à-dire, x_{n+1} ne dépend que de x_n . Si il dépend de plus de termes précédents, la méthode est dite **multi-pas** : $x_{n+1} = x_n + h \cdot \phi(t_n, x_n, \dots, x_{n-k}, h)$. Et pour calculer le ϕ , il y a plusieurs deux cas. Pour se simplifier l'écriture, on suppose la méthode à un pas, mais ça se généralise. Il se peut qu'on puisse calculer $\phi(t_n, x_n, h)$ directement (on a une formule explicite), auquel cas la méthode est **explicite**. Mais parfois, on doit calculer $\phi(t_n, x_n, h)$ en résolvant une équation, auquel cas la méthode est **implicite** (on peut alors écrire $x_{n+1} = x_n + h \cdot \tilde{\phi}(t_n, x_n, t_{n+1}, x_{n+1}, h)$, de manière à suggérer la forme implicite du problème).

→ Méthodes : exemples.

On a plein de manières d'approximer une solution, autant choisir la plus pratique (efficace ou simple à appliquer).

■ La motivation.

D - Donc, comment tu approximerais la solution à l'équation différentielle $x' = x$ ($x(0) = 1$) ?

T - Par exemple, une méthode d'Euler explicite va approximer $x(t+h)$ par $x(t) + hx'(t)$, donc $x(nh)$ par $(1+h)^n$ par récurrence.

D - Donc, avec un pas de t/n , on approxime e^t par $(1 + \frac{t}{n})^n$!

T - Oui, c'est une forme bien connue, et ça converge vers la bonne valeur quand $h \rightarrow 0$ (c'est-à-dire $n \rightarrow +\infty$).

■ Construction d'une méthode.

Dans cette section, on résout des équations autonomes (de la forme $x'(t) = f(x(t))$), et on construit des méthodes à un pas, avec un pas h constant. Pour ça, on peut réécrire $x(t_{n+1}) = \int_{t_n}^{t_{n+1}} x'(t) dt = \int_{t_n}^{t_{n+1}} f(t, x(t)) dt$. On ne peut toujours pas calculer cette intégrale à la main, mais on peut l'approximer par une des nombreuses méthodes de quadrature qu'on connaît.

■ Méthode d'Euler explicite.

On utilise la méthode des rectangles à gauche. Donc l'intégrale est approximée par $(t_{n+1} - t_n) \cdot f(t_n, x(t_n)) = hf(t_n, x_n)$. Ainsi obtient-on une méthode explicite, dite méthode d'Euler explicite :

$$\begin{cases} x_{n+1} = x_n + hf(t_n, x_n) \\ x_0 \text{ fixé dans } \mathbb{R}^m \end{cases}$$

(C'est-à-dire, $\phi(t_n, x_n, h) = f(t_n, x_n)$.)

■ Méthode d'Euler implicite.

On utilise la méthode des rectangles à droite. Donc l'intégrale est approximée par $(t_{n+1} - t_n) \cdot f(t_{n+1}, x(t_{n+1})) = hf(t_{n+1}, x_{n+1})$. Ainsi obtient-on une méthode implicite, dite méthode d'Euler implicite :

$$\begin{cases} x_{n+1} = x_n + hf(t_{n+1}, x_{n+1}) \\ x_0 \text{ fixé dans } \mathbb{R}^m \end{cases}$$

(C'est-à-dire, $\phi(t_{n+1}, x_{n+1}, h) = f(t_{n+1}, x_{n+1})$.)

■ Méthode de Crank-Nicolson.

On utilise la méthode des trapèzes. Donc l'intégrale est approximée par $(t_{n+1} - t_n) \cdot \frac{f(t_n, x(t_n)) + f(t_{n+1}, x(t_{n+1}))}{2} = h \frac{f(t_n, x_n) + f(t_{n+1}, x_{n+1})}{2}$. Ainsi obtient-on une méthode implicite (voire semi-implicite pour certains), dite méthode de Crank-Nicolson :

$$\begin{cases} x_{n+1} = x_n + h \frac{f(t_{n+1}, x_{n+1}) + f(t_n, x_n)}{2} \\ x_0 \text{ fixé dans } \mathbb{R}^m \end{cases}$$

(C'est-à-dire, $\phi(t_{n+1}, x_{n+1}, h) = \frac{f(t_{n+1}, x_{n+1}) + f(t_n, x_n)}{2}$.)

■ Méthode de Heun.

Pour la méthode de Heun, on part d'un Crank-Nicolson, et on approxime le x_{n+1} par une méthode d'Euler explicite. Au final, on obtient une méthode explicite.

$$\begin{cases} x_{n+1} = x_n + \frac{h}{2}[f(t_n, x_n) + f(t_{n+1}, x_n + hf(t_n, x_n))] \\ x_0 \text{ fixé dans } \mathbb{R}^m \end{cases}$$

i.e. $\phi(t_n, x_n, h) = [f(t_n, x_n) + f(t_{n+1}, x_n + hf(t_n, x_n))]/2$

■ Méthodes de Runge-Kutta.

Sur la même idée, on pourrait partir de la quadrature donnée par une méthode du point milieu plutôt que d'une méthode des trapèzes, et expliciter le terme $x_{n+\frac{1}{2}}$. Le résultat, explicite, est appelé méthode de Runge-Kutta d'ordre 2 (RK2 en bref).

$$\begin{cases} x_{n+1} = x_n + hf(t_n + \frac{h}{2}, x_n + \frac{h}{2}f(t_n, x_n)) \\ x_0 \text{ fixé dans } \mathbb{R}^m \end{cases}$$

(C'est-à-dire, $\phi(t_n, x_n, h) = f(t_n + \frac{h}{2}, x_n + \frac{h}{2}f(t_n, x_n))$.) Traditionnellement, on aime bien écrire ce calcul comme :

$$k_1 \leftarrow f(t_n, x_n), k_2 \leftarrow f(t_n + \frac{h}{2}, x_n + \frac{h}{2}k_1), x_{n+1} \leftarrow x_n + hk_2$$

On peut généraliser l'idée, en construisant des méthodes de la forme :

$$k_1 \leftarrow f(t_n, x_n), k_2 \leftarrow f(t_n + c_2h, x_n + (a_{21}k_1)h), \dots$$

$$k_s \leftarrow f(t_n + c_sh, x_n + (a_{s1}k_1 + \dots + a_{s,s-1}k_{s-1})h)$$

$$x_{n+1} \leftarrow x_n + h(b_1k_1 + \dots + b_sk_s)$$

Avec b_i , a_{ij} et c_j bien choisis; souvent, on les indique dans des tableaux de Butcher :

0				
c_2	a_{21}			
\vdots	\vdots	\ddots		
c_s	a_{s1}	\dots	$a_{s,s-1}$	
	b_1	\dots	b_{s-1}	b_s

La méthode de Runge-Kutta (classique) est celle pour $s=4$, de manière explicite :

$$k_1 \leftarrow f(t_n, x_n), k_2 \leftarrow f(t_n + h/2, x_n + hk_1/2)$$

$$k_3 \leftarrow f(t_n + h/2, x_n + hk_2/2), k_4 \leftarrow f(t_n + h, x_n + hk_3)$$

$$x_{n+1} \leftarrow x_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

→ Résolution d'équations.

Si on veut mettre en place une méthode implicite, il faut pouvoir trouver (algorithmiquement) un point fixe.

■ La motivation.

D - Donc pour résoudre $(x'(t), y'(t)) = (y(t), -x(t))$ avec un Euler implicite, il faut trouver une solution à $(x_{n+1}, y_{n+1}) = (x_n + hy_{n+1}, y_n - hx_{n+1})$.

T - C'est un système, on pourrait le résoudre à la main. Mais ...

D - ... numériquement, il y a un moyen ? On peut le ramener à une équation de la forme $H(x_{n+1}, y_{n+1}) = 0$, mais après ?

T - Si H ne s'annule pas en un point v , et que la dérivée non plus, on peut aller dans une direction qui rapproche la valeur de H vers 0 ; on remplace l'approximation v d'une racine, par $v - (DH(v))^{-1}H(v)$. C'est l'idée de la méthode de Newton.

■ Méthode de point fixe.

L'algorithme le plus simple est de construire $x_{n+1} = \phi(x_n)$ avec x_0 fixé et $\phi : E \rightarrow E$ ($E = \mathbb{R}^m$) la fonction dont on cherche les points fixes. Si la suite ainsi construite converge, alors la limite est bien un point fixe. Pour garantir cette convergence, on peut appliquer le **théorème du point fixe de Banach** : si ϕ est strictement contractante, alors ϕ admet un unique point fixe \bar{x} et quelque soit le x_0 choisi, $(x_n)_n$ converge vers \bar{x} .

■ Généralisation.

La stricte contractance est une condition forte ; un théorème plus applicable est : Si $\phi : \mathcal{U} \rightarrow \mathcal{U}$ (avec \mathcal{U} ouvert de E) est \mathcal{C}^1 et admet un point fixe $\bar{x} \in \mathcal{U}$, alors avoir $\rho(\mathcal{J}_\phi(\bar{x})) < 1$ nous garantit l'existence d'un voisinage $B(\bar{x}, \delta)$ de \bar{x} tel que pour tout $x_0 \in B(\bar{x}, \delta)$, alors la suite donnée par la méthode converge vers \bar{x} . Comme $\rho(A) \leq \|A\|$ pour toute norme matricielle, le théorème s'applique aussi si $\|A\| < 1$.

■ Ordre de convergence.

On s'intéresse à des méthodes qui construisent une suite convergeant vers la valeur qu'on cherche, comme la méthode du point fixe. Une telle méthode est dite **convergente à l'ordre p** exactement si $\frac{\|x_{n+1} - \bar{x}\|}{\|x_n - \bar{x}\|^p}$ est une suite convergeant vers un réel non-nul C ; la vitesse de convergence est alors C . Plus généralement, une méthode converge au moins à l'ordre p , si il existe $K > 0$ tel que $\|x_{n+1} - \bar{x}\| \leq K \|x_n - \bar{x}\|^p$ pour tout n .

■ Théorème de convergence.

On peut donner l'ordre pour la méthode du point fixe, en fonction de la norme (matricielle) de la Jacobienne de ϕ . On se place sur $\Omega \in \mathbb{R}^m$ un ouvert connexe, avec $\phi : \Omega \rightarrow \mathbb{R}^m$ une fonction \mathcal{C}^2 , de différentielle seconde bornée sur Ω . Alors, si on pose la suite $\begin{cases} x_{n+1} = \phi(x_n) \\ x_0 \in \Omega \end{cases}$ qui converge vers \bar{x} avec $x_n \neq \bar{x}$, on peut dire que : - si $0 < \|\mathcal{J}_\phi(\bar{x})\| < 1$, alors la convergence est d'ordre 1 exactement, et de vitesse $\|\mathcal{J}_\phi(\bar{x})\|$ - si $\|\mathcal{J}_\phi(\bar{x})\| = 0$, alors la convergence est au moins d'ordre 2, et il existe un $C > 0$ tel que $\|x_{n+1} - \bar{x}\| \leq C \|x_n - \bar{x}\|^2$.

■ Méthode de Newton.

Voici une autre méthode de résolution d'équations. On veut donc résoudre $H(\bar{x}) = 0$. On supposera que $H : \mathbb{R}^m \rightarrow \mathbb{R}^m$ est \mathcal{C}^2 . On pose alors

$$\begin{cases} x_{n+1} = x_n - (DH(x_n))^{-1}H(x_n) \\ x_0 \in \mathbb{R} \end{cases}$$

C'est la **méthode de Newton(-Raphson)**. On notera que ça revient à appliquer une méthode de point fixe à la fonction $\phi(x) = x - (DH(x))^{-1}H(x)$, qui admet les \bar{x} (avec $H(\bar{x}) = 0$) comme points fixes.

Alors si il existe $\bar{x} \in \mathbb{R}^m$ avec $H(\bar{x}) = 0$ et $\mathcal{J}_H(\bar{x})$ inversible, alors $(x_n)_n$ est bien-définie, et la méthode converge localement à l'ordre 2 ; où localement signifie qu'il y a un voisinage $B(\bar{x}, \delta)$ de \bar{x} tel que x_0 pris dedans nous donne $x_n \rightarrow \bar{x}$.

Dans la pratique, nous devons arrêter l'algorithme pour en récupérer un résultat. Parmi les critères d'arrêts possibles, les plus courants sont : un nombre maximum d'itération N_{iter} , une condition d'erreur locale ("tant que $\|x_{n+1} - x_n\| > \epsilon_{erreur}$ ") ou une condition d'erreur globale ("tant que $\|H(x_n)\| > \epsilon_{erreur}$ "). On peut aussi faire une combinaison de ces critères.

■ Exemple.

Commençons par illustrer une méthode de point fixe. Si on applique une méthode Euler implicite à l'équation différentielle $x'(t) = x(t)$ (avec $x(0) = (1, 0)$), il faut résoudre $x_{n+1} = x_n + h \cdot x_{n+1}$, donc trouver un point fixe de $\phi(y) = x_n + hy$. On a $\|\phi(y) - \phi(z)\| = h\|y - z\|$, donc ϕ h -Lipschitzienne. Donc si $h < 1$, ϕ est strictement contractante et on peut appliquer la méthode pour calculer x_{n+1} . Note : on peut aussi calculer x_{n+1} directement dans ce cas très particulier.

Maintenant, mettons en œuvre une méthode de Newton. Il faut donc résoudre $x_{n+1} = x_n + h \cdot x_{n+1}$, donc $x_{n+1} - (x_n + h \cdot x_{n+1}) = 0$; on pose alors $H(y) = y - (x_n + h \cdot y)$. H est bien \mathcal{C}^2 , $H(\bar{y}) = 0$ (ce \bar{y} existe bien, en l'occurrence c'est $\frac{x_n}{1-h}$), et $\mathcal{J}_H(\bar{y}) = \begin{pmatrix} 1-h & 0 \\ 0 & 1-h \end{pmatrix}$; donc comme cette matrice est non-nulle si $h \neq 1$, on peut utiliser la méthode de Newton, et elle fonctionnera pourvu que le y_0 soit assez proche de \bar{y} . Ici, on peut prendre $y_0 = x_n$ sans problème.

→ Convergence.

On obtient une approximation oui, mais est-elle assez proche de la réalité ?

■ La motivation.

D – Si j'utilise une méthode d'Euler explicite sur $x'(t) = -100x(t)$, alors les solutions approchées ressembleront à $x_n = (1 - 100h)^n$. Sauf que si h est trop grand, ça n'a pas la bonne allure !

T – Oui, il faut $0 < h < 1/50$ pour que ça marche. Il existe des problèmes en physique où il faut prendre h très petit ; ce sont des problèmes raides.

D – Mais est-ce qu'il y a toujours un h assez petit ?

T – Que si la méthode est dite "convergente" ; et pour vérifier ça, il y a deux critères : stable, consistance à un ordre $p \geq 1$.

■ Erreur de consistance (troncature).

Notons $\Phi(t_n, x, h)$ la fonction qui approxime x_{n+1} suivant une méthode à partir de la valeur exacte $x(t_n)$. Pour un Euler Explicite, ce serait

$$\Phi(t_n, x, h) = x(t_n) + h \cdot f(t_n, x(t_n))$$

en général, c'est de la forme $\Phi(t_n, x, h) = x(t_n) + h \cdot \phi(t_n, x, h)$. On peut alors construire l'**erreur de consistance (de troncature)** comme l'erreur entre l'approximation et la solution exacte,

$$\epsilon(t_n, x, h) := x(t_n + h) - \Phi(t_n, x, h)$$

Si $\epsilon(t_n, x, h)$ est $O(h^{p+1})$, on dit la méthode est **consistante à l'ordre p**. Note : certaines sources considèrent plutôt la quantité $\sum_{n=0}^{N-1} |\epsilon(t_n, x, h)|$ comme erreur de consistance "globale", comparée à $\epsilon(t_n, x, h)$ l'erreur locale.

Exemple : pour une méthode d'Euler explicite,

$$\epsilon_n := \epsilon(t_n, x, h) = x(t_n + h) - (x(t_n) + hf(t_n, x(t_n)))$$

Par un développement limité de x (qu'on suppose \mathcal{C}^2), on a

$$\epsilon_n = x(t_n) + hx'(t_n) + O(h^2) - x(t_n) - hf(t_n, x(t_n)) = O(h^2)$$

(puisque $x'(t_n) = f(t_n, x(t_n))$). Donc Euler explicite est consistant à l'ordre 1.

■ Stabilité.

La notion de stabilité permet de contrôler l'erreur obtenue si on applique la méthode sur un ordinateur (qui va donc arrondir les valeurs à cause d'une précision finie). Pour ça, on pose deux suites : x_n défini par $x_{n+1} = x_n + h \cdot \phi(t_n, x_n, h)$ et $y_{n+1} = y_n + h \cdot \phi(t_n, y_n, h) + \alpha_n$ la version perturbée par une suite $(\alpha_n)_n$ d'erreurs d'arrondi. Pour avoir la **stabilité**, il faut avoir une constante $K > 0$ et un h assez petit afin que

$$\|x_n - y_n\| \leq K \cdot \left(\|x_0 - y_0\| + \sum_{k=0}^{N_h-1} \|\alpha_k\| \right)$$

■ Stabilité absolue.

La **stabilité absolue** concerne le comportement sur un temps long, c'est-à-dire sans borne de temps. Précisément, un schéma est absolument stable si pour $h > 0$, l'approximation x_n reste dans une boule $B(0, R)$ (pour un $R > 0$), et donc est bornée (ainsi que sa limite si elle existe). Pour l'exemple de la motivation, si $h > 0.02$, la méthode n'est pas absolument stable puisque la solution tend vers l'infini.

■ Lemme de Gronwall discret.

Soit $a > 0$ et deux suites $(A_n)_n$ et $(B_n)_n$, si $A_{n+1} \leq a \cdot A_n + B_n$, alors

$$A_n \leq a^n A_0 + \sum_{i=0}^{n-1} a^{n-(i+1)} B_i$$

et en posant $u = a - 1$ on a

$$A_n \leq e^{nu} A_0 + \sum_{i=0}^{n-1} e^{(n-(i+1))u} B_i$$

Ce lemme a une version continue. Mais surtout, il nous permet de garantir que si $\phi(t_n, x_n, h)$ est continue et L -lipschitzienne en sa deuxième variable, alors la méthode est stable.

■ Convergence.

On a alors le théorème suivant : si une méthode est consistante à l'ordre p , et stable, alors elle est **convergente à l'ordre p**, dans le sens où il existe un $C > 0$ tel que pour h assez petit, $\max_{0 \leq n \leq N_h} \|x(t_n) - x_n\| \leq Ch^p$. Cela veut dire qu'en prenant h assez petit, on peut baisser l'erreur autant qu'on le souhaite. Donc quand $h \rightarrow 0$, l'approximation "converge" vers la solution.