

Introduction aux Méthodes Numériques pour les équations différentielles ordinaires

Math 603, L3, Université Savoie-MtBlanc

KRZYSZTOF WORYTKIEWICZ

1 CRITIQUE DE LA RAISON ANALYTIQUE

1.1 DISCOURS DE LA MÉTHODE

Analysons à titre d'exemple la méthode dite *variation de la constante* qui marche pour des équas diff de type

$$y' = a(t)y + b(t) \quad (*)$$

Afin de trouver une solution on procède comme suit (*recette de cuisine*):

1. L'équa *homogène*

$$y' = a(t)y$$

a pour solutions

$$y_C(t) = Ce^{A(t)}$$

avec $C \in \mathbb{R}$ arbitraire et $A(t) = \int a(t)dt$. Mais une *formule* pour la primitive $A(t)$ peut ne pas exister (ou pire: la primitive peut ne pas exister...).

2. Soit y_p une *solution particulière* de (*). L'ensemble des solutions de (*) est

$$\{y_p + y_C | C \in \mathbb{R}\}$$

Mais comment trouve-t-on une y_p ??

1. On choisit une solution de l'équa homogène qui convient, typiquement

$$y_1(t) = e^{A(t)} \quad (**)$$

(donc on pose $C := 1$);

2. On *mise* sur une solution particulière de la forme

$$y_p(t) = c(t)y_1(t) \quad (***)$$

et on cherche à déterminer la fonction c . Existe-t-elle? Si oui peut-on l'exprimer au moyen d'une *formule* ??

3. Il s'avère que oui, peut-être: sous les hypothèses ci-dessus on peut montrer que

$$c(t) = \int b(t)e^{-A(t)}dt$$

avec $A(t)$ comme ci-dessus. Même problème donc..

1.2 FORMULES

On sait depuis les Babyloniens (c'est un fait historique) que les solutions d'une équa polynomiale de 2nd degré

$$ax^2 + bx + c = 0$$

sont données par la *formule*

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Ici la *formule* est formée au moyen de puissances, racines et fractions. En algèbre on appelle cela *solution par radicaux*. De telles solutions par radicaux existent également en toute généralité pour les équations polynomiales de 3e et 4e degré. Par contre à partir de 5e degré ce n'est plus le cas en général, le fameux *théorème de Galois* (1830) donne un critère précis pour l'existence d'une solution par radicaux (le groupe des permutations des racines est dit *résoluble*).

De telles situations se rencontrent dans divers domaines de maths, elles ne sont pas exclusives à l'algèbre. Ce que l'on entend par *formule* est largement une convention à base de bon sens. Par exemple on ne considère pas l'opération de *somme infinie*. Dans le contexte des équations diff, une solution exprimable par une telle formule s'appelle *solution analytique*. Elle est formée au moyen de

- constantes (réels ou complexes);
- variables;
- additions, multiplications etc;
- sommes finies et produits finis;
- fractions continues finies;
- puissances (en particulier racines);
- logarithmes;
- fonctions trigo et hyperboliques ainsi que leurs inverses.

On a vu que la méthode dite *variation de la constante* est basée sur l'existence de primitives. Mais ces primitives, même si elles existent, sont-elles toujours exprimables au moyen de formules? Il s'avère que ce n'est pas toujours le cas. Le contre-exemple classique: on a

$$\int e^{-x^2} dx = \frac{2}{\pi} \int_0^x e^{-t^2} dt$$

à constante près. L'intégrale n'admet pas de solution analytique, bien qu'elle définit uniquement une fonction (notée traditionnellement erf).

Résumons: pour la résolution d'une équation diff on rencontre essentiellement 2 problèmes:

1. Une solution existe, mais on ne connaît pas de solution analytique (c'est le cas général);
2. Une solution analytique existe, mais elle n'est pas « gérable » (par exemple la formule prend 50 pages...).

La bonne nouvelle: il existe des *critères d'existence de solutions*. Une fois l'existence établie, on peut *l'approximer numériquement* si une solution analytique pratique n'est pas disponible.

2 PROBLÈMES DE CAUCHY

DÉFINITION. Un *problème de Cauchy* est une équation diff de 1er ordre de la forme

$$y' = f(t, y)$$

subjecte à une *condition initiale*

$$y(t_0) = y_0$$

En fait de Cauchy-Lipschitz-Picard-Lindelöf



(ai-je oublié quelqu'un?..).

Un problème de Cauchy admet-t-il toujours une solution? Une solution unique?

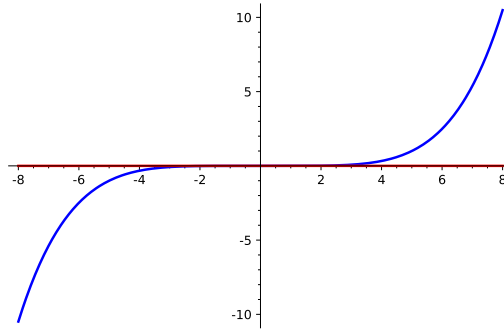
Exemple. Considérons le problème de Cauchy

$$\begin{aligned} y' &= y^{\frac{4}{5}} \\ y(0) &= 0 \end{aligned}$$

On voit par inspection que

$$\begin{aligned} y(t) &= \left(\frac{t}{5}\right)^5 \\ y(y) &= 0 \end{aligned}$$

sont des solutions distinctes:



THÉORÈME. (CLPL+) Soient

- $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ un ouvert et $f: \Omega \rightarrow \mathbb{R}$ continue;
- $(t_0, y_0) \in \Omega$;
- $\mathcal{R} = \{(t, y) \in \Omega \mid |t - t_0| \leq a, \|y - y_0\| \leq b\}$ un compact.

Supposons qu'il existe des constantes

- i. $L \in \mathbb{R}_+$ telle que $|f(t, u) - f(t, v)| \leq L\|u - v\|$ pour tout $(t, u), (t, v) \in \mathcal{R}$;
- ii. $M \in \mathbb{R}_+$ telle que $|f(t, u)| \leq M$

Posons $T := \min\left(a, \frac{b}{M}\right)$ et $I := [t_0 - T, t_0 + T]$. Le problème de Cauchy

$$\begin{aligned} y' &= f(t, y) \\ y(t_0) &= y_0 \end{aligned}$$

admet une unique solution $y: I \rightarrow X$, maximale de classe C^1 .

C'est une des variantes possibles. Observons que l'hypothèse (i) dit que f est *Lipschitz en deuxième variable* (on cale t) sur \mathcal{R} , L étant la *constante de Lipschitz*. On démontre le théorème en définissant par récurrence une suite de fonctions

- i. $y_0(t) := y_0$;
- ii. $y_n(t) := y_0 + \int_{t_0}^t f(\xi, y_{n-1}(\xi)) d\xi$

puis en montrant que cette suite converge uniformément vers la fonction

$$y(t) = y_0 + \int_{t_0}^t f(\xi, y(\xi)) d\xi$$

(convergence par rapport à une norme dans un espace de fonctions). Celle-ci est une solution du pb de Cauchy, son unicité est une conséquence de la condi de Lipschitz. Les détails de la preuve sont loin d'être triviaux.

3 MÉTHODES NUMÉRIQUES CLASSIQUES

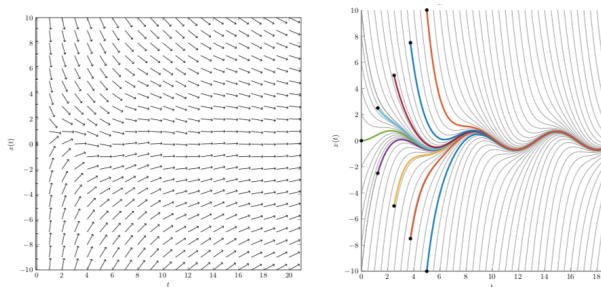
En pratique, l'itération CLPL+ est trop coûteuse pour être utilisée explicitement. Reprenons l'équa diff faisant partie d'un problème de Cauchy

$$y' = f(t, y) \quad (*)$$

par exemple

$$f(t, y) = -y + \sin(t)$$

de tout à l'heure. On peut interpréter *géométriquement* (*) comme la donnée d'un champ de vecteurs



On calcule en chaque point $P = (t, y)$ le vecteur directeur de la droite de pente $f(t, y)$ passant par P . Si (*) vérifie les hypothèses CLPL+, toute condi initiale donne une courbe qui en tout point aura un vecteur tangent appartenant au champ. Ce point de vue géométrique est le point de départ des méthodes numériques classiques Euler et Runge-Kutta, respectivement.

3.1 EULER A ENCORE FRAPPÉ..



Parmi les méthodes classiques c'est la plus simple: on se donne un problème de Cauchy

$$\begin{aligned} y' &= f(t, y) \\ y(t_0) &= y_0 \end{aligned}$$

qui admet une solution dans l'intervalle $I := [t_0, T]$ (typiquement en vertu de Cauchy-Lipschitz). On discrétise I avec un *pas de maillage*. On se donne une précision $N > 0$ et on pose

$$\begin{aligned} h &:= \frac{T - t_0}{N} \\ t_n &:= t_0 + nh \end{aligned}$$

puis on approxime la solution avec la suite

$$(\hat{y}_n)_{0 \leq n \leq N}$$

qu'on calcule par récurrence:

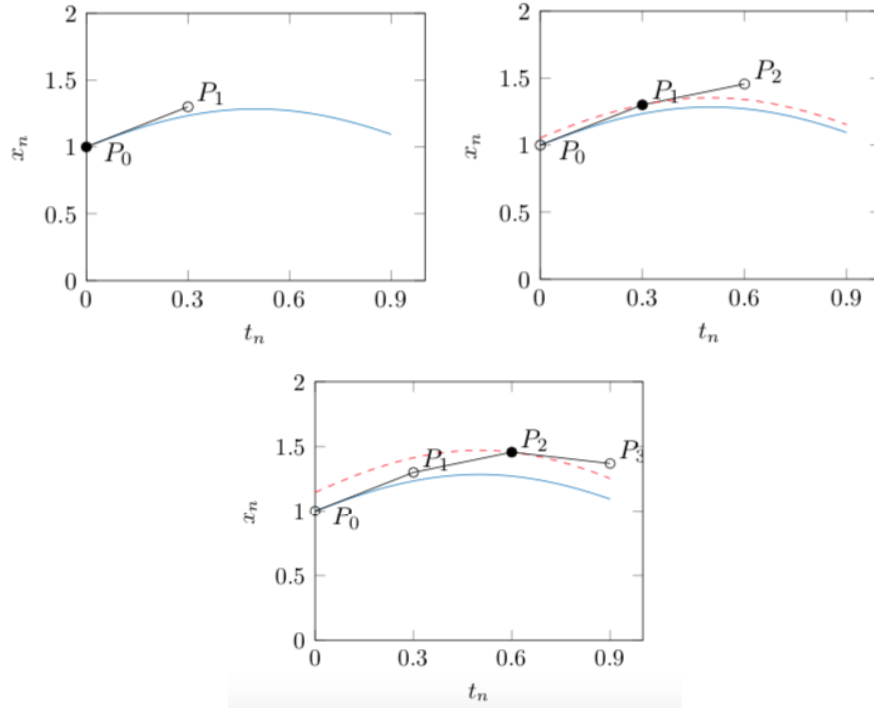
$$\begin{aligned} \hat{y}_0 &= y_0 \\ \widehat{y_{n+1}} &= \hat{y}_n + f(t_n, \hat{y}_n)h \end{aligned}$$

cette à dire en « suivant » les segments de tangentes.

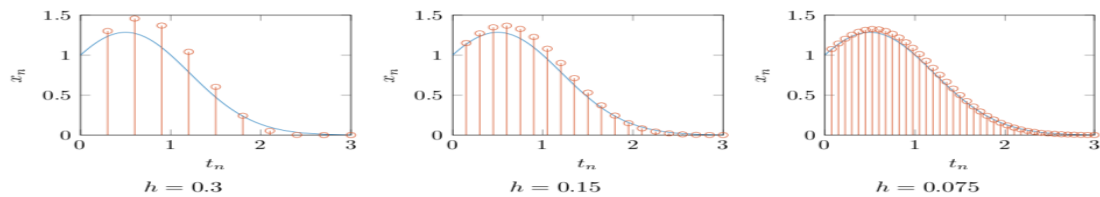
Exemple. On se donne le problème de Cauchy

$$\begin{aligned} y' &= (1-2t)y \\ y(t_0) &= 1 \end{aligned}$$

que l'on aimerait approximer dans l'intervalle $I := [0, 0.9]$. Traçons les 3 premières itérations pour $h = 0.3$:



Hmm, pas très satisfaisant. On peut par contre observer que plus le pas de maillage est petit, plus la précision est grande:



3.2 LE PROBLÈME AVEC EULER

DÉFINITION. Une méthode numérique à 1 pas d'un PC admettant une solution est une itération

$$\begin{aligned} \hat{y}_0 &= y_0 \\ \widehat{y_{n+1}} &= \Psi_h(t_n, \hat{y}_n) \end{aligned}$$

L'erreur globale de la méthode est la quantité

$$E_\Psi \stackrel{\text{def.}}{=} \hat{y}_N - y(t_N)$$

où y est la solution du PC.

Il est en général impossible de déterminer l'erreur globale de manière exacte. Les théorèmes de précision sont typiquement des résultats en termes d'approximations asymptotiques (la notation de Landau...).

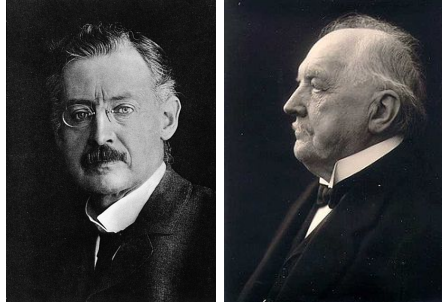
THÉORÈME. Pour la méthode d'Euler on a $E = \mathcal{O}(h)$

Cela veut dire que si l'on souhaite une marge d'erreur de 10^{-8} il faut un pas de maillage

$$h = 10^{-8}$$

Cela implique que si l'on itère sur $I = [0, 1]$, il faut 10^8 pas et donc 10^8 évaluation de fonctions. Ce n'est donc pas très satisfaisant du point de vue de la précision. Par contre la simplicité de la méthode d'Euler fait qu'elle est toujours utilisée dans les situations où le temps de calcul est important, ce qui est par exemple le cas pour les *systèmes embarqués*.

3.3 LA FAMILLE RUNGE-KUTTA



Cette famille de méthodes numériques améliore Euler en précision en « subdivisant les subdivisions »

$$\begin{aligned} k_1 &:= f(t_n, y_n) \\ k_2 &:= f(t_n + c_2 h, y_n + h a_{2,1} k_1) \\ k_3 &:= f(t_n + c_3 h, y_n + h(a_{3,1} k_1 + a_{3,2} k_2)) \\ k_4 &:= f(t_n + c_4 h, y_n + h(a_{4,1} k_1 + a_{4,2} k_2 + a_{4,3} k_3)) \\ &\vdots \\ k_s &:= f\left(t_n + c_s h, y_n + h \sum_{j=1}^{s-1} a_{s,j} k_j\right) \\ y_{n+1} &:= y_n + h \sum_{i=1}^s b_i k_i \end{aligned}$$

Les k_i 's font partie du champ de tangentes donné par le problème de Cauchy. Un pas d'approximation consiste à déterminer le prochain data point au moyen de combinaisons linéaires des tangentes k_i . Toute instance de la méthode Runge-Kutta est donc déterminée par les coefficients $a_{i,j}$, b_j et c_i . Ces coefficients s'organisent en un *Tableau de Butcher*

0	$\left \begin{array}{cccc} a_{2,1} & & & \\ a_{3,1} & a_{3,2} & & \\ \vdots & & \ddots & \\ a_{s,1} & a_{s,2} & \cdots & a_{s,s-1} \end{array} \right.$				
c_2					
c_3					
\vdots					
c_s					
<hr/>					
	b_1	b_2	\cdots	b_{s-1}	b_s

On requiert

$$\sum_{j=1}^{i-1} a_{i,j} = c_i$$

pour $2 \leq i \leq s$ pour des raisons de *consistence*.

3.4 LA CLASSIQUE RK4

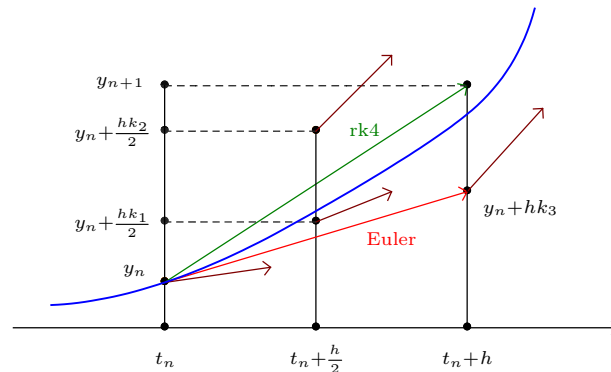
RK4 est donnée par le tableau

0				
$\frac{1}{2}$	$\frac{1}{2}$			
$\frac{1}{2}$	0	$\frac{1}{2}$		
1	0	0	1	
<hr/>				
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

c'est à dire

$$\begin{aligned}
 k_1 &:= f(t_n, y_n) \\
 k_2 &:= f(t_n + h/2, y_n + hk_1/2) \\
 k_3 &:= f(t_n + h/2, y_n + hk_2/2) \\
 k_4 &:= f(t_n + h, y_n + hk_3) \\
 y_{n+1} &:= y_n + h(k_1 + 2k_2 + 2k_3 + k_4)/6
 \end{aligned}$$

En général le procédé améliore les simples tangentes d'Euler:



THÉORÈME. L'erreur globale de RK4 est $E = \mathcal{O}(h^4)$.

Cela veut dire que si l'on souhaite une marge d'erreur de 10^{-8} il faut un pas de maillage

$$h = 10^{-2}$$

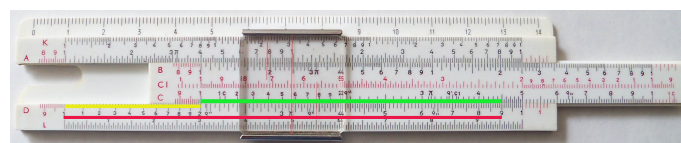
Cela implique que si l'on itère sur $I = [0, 1]$, il faut 100 pas et 400 évaluation de fonctions. Si l'on choisit un pas de maillage

$$h = 10^{-3}$$

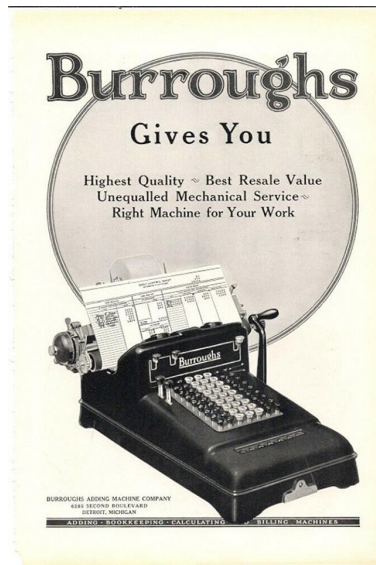
(c'est à dire 1000 pas et 4000 évaluation de fonctions) on obtient une marge d'erreur de 10^{-12} . Les méthodes Runge-Kutta, en particulier RK4, sont parmi les plus utilisées encore de nos jours.

4 MACHINES ET REPRÉSENTATION DES FLOTTANTS

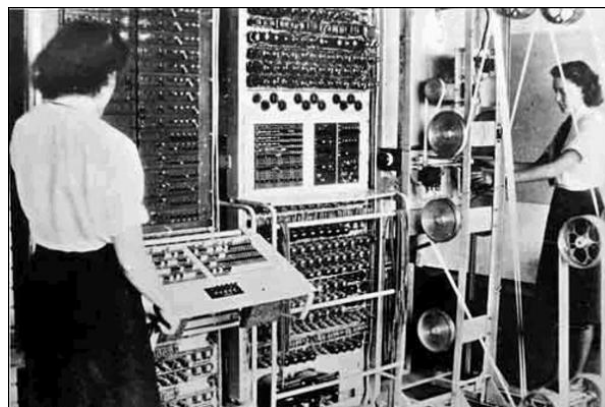
Jusqu'ici tout a l'air idyllique, mais... Il faut en premier lieu interpréter correctement et placer dans leur contexte les résultats de précision ci-dessus. Les instruments de calcul à l'époque de Runge et de Kutta étaient analogiques et s'appelaient *règles à calcul*:



Les résultats sur la précision RK sont donc valides (si l'on choisit un $\varepsilon > 0$ adéquat...). Mais l'avènement des machines à calcul digitales a changé la donne. Au début électromécaniques



puis électroniques à lampes (ici le Colossus II)



et finalement électroniques à transistors sur circuits intégrés de plus en plus miniaturisés (ici l'IBM 360)



vous connaissez la suite... Toutes ces machines partagent une même caractéristique: un nombre limité de bits pour représenter les nombres. Cela reste largement vrai aujourd'hui. Pour ce qui est de la représentation des nombres réels dits « flottants » dans ce contexte, le standard IEEE 754 introduit en 1985 reste la base

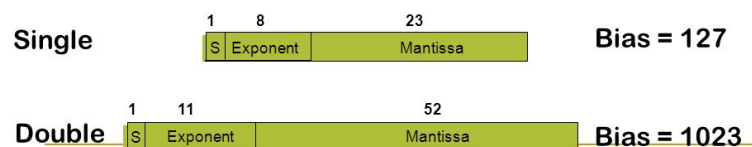
Floating point numbers– the IEEE standard

■ IEEE Standard 754

- Most (but not all) computer manufactures use IEEE-754 format
- Number represented:

$$(-1)^S * (1.M) * 2^{(E - \text{Bias})}$$

2 main formats: single and double



Il s'agit d'un *choix*: le format IEEE est un compromis entre coûts de production, efficacité de calcul et sa précision. La quasi-totalité des processeurs sur le marché fonctionnent avec ce format. Mais les conséquences ne sont pas anodines, parmi elles:

- l'ensemble des nombres représentés est un sous-ensemble *non-dense* de \mathbb{Q} ;
- l'addition *n'est pas* associative;
- la précision est *au mieux* 15 décimales correctes pour le format *double* et 32 pour le format *quad* (64 bit).

Les méthodes numériques sont étudiées depuis longtemps, il existe une multitude d'astuces permettant de maîtriser les calculs en virgule flottante. Par contre, la limite de précision est « hard ».

5 PRÉCISION ARBITRAIRE

Parfois il faut plus que 32 décimales, par exemple pour des calculs en rapport avec:

- la stabilité à long terme du système solaire;
- certains problèmes de la mécanique des fluides (*l'attracteur de Lorentz*);
- certains problèmes de la physique des sous-particules (*Chromodynamique quantique*);
- la cryptographie quantique;
- la géométrie algorithmique;
- votre problème favori à précision gigantesque...

Est-ce possible de calculer en *précision arbitraire*, au sens que la seule limite de précision est imposée par la mémoire disponible? Dans 80's quand a vu le jour le format IEEE, la précision arbitraire était impensable en premier lieu vu le matériel de l'époque, mais aussi au niveau des algorithmes connus. Ceci a bien changé depuis car le matériel et l'algorithmique ont progressé, considérablement. De nos jours la précision arbitraire est tout à fait envisageable bien que le procédé reste coûteux. Par contre cela se fait essentiellement au moyen d'une autre approche, basée sur l'expansion de Taylor. Le dilemme classique:

- A) Runge-Kutta marche mieux en virgule flottante, mais ce n'est plus le cas en précision arbitraire.

B) Taylor est une approche naturelle pour la précision arbitraire, mais nécessite les dérivées d'ordre supérieur, difficiles à approximer numériquement.

Mais B) n'est plus un problème lorsqu'on a une capacité de calcul symbolique efficace. Alors comment ça marche? Si l'on se donne un problème de Cauchy *on ne connaît pas* la solution $y(t)$. Comment construire alors son expansion de Taylor?

Soit $g \in C^\infty$. Rappelons que le *polynôme de Taylor d'ordre n* de g autour de t_0 est donné par

$$\begin{aligned} G_n(t) &= \sum_{i=0}^n \frac{1}{i!} g^{(i)}(t_0) \\ &= g(t_0) + g'(t_0)(t - t_0) + \frac{1}{2}g''(t_0)(t - t_0) + \frac{1}{6}g'''(t_0)(t - t_0) + \cdots + \frac{1}{n!}g^{(n)}(t - t_0) \end{aligned}$$

et on a

$$g(x) = \lim_{n \rightarrow \infty} G_n(x)$$

Supposons un problème de Cauchy

$$\begin{aligned} y' &= f(t, y) \\ y(t_0) &= y_0 \end{aligned}$$

avec $f \in C^\infty$ et $y(t)$ une solution. On ne connaît bien sûr pas explicitement cette dernière (puisque on cherche à l'approximer...), par contre le problème de Cauchy fournit suffisamment d'information pour *l'évaluer en t_0* ainsi que toutes ses dérivées. On a gratis les 2 premiers termes du n -ième polynôme de Taylor

$$\begin{aligned} Y_n(t) &= y(t_0) + y'(t_0)(t - t_0) + \sum_{i=2}^n \frac{1}{i!} y^{(i)}(t - t_0) \\ &= y_0 + f(t_0, y_0)(t - t_0) + \sum_{i=2}^n \frac{1}{i!} y^{(i)}(t - t_0) \end{aligned}$$

car

$$\begin{aligned} y'(t_0) &= f(t_0, y(t_0)) \\ &= f(t_0, y_0) \end{aligned}$$

Pour continuer, on a besoin des dérivées d'ordre supérieur.

THÉORÈME. (DÉRIVÉE COMPOSÉE À 2 VARIABLES) Soit $u(x(t), y(t))$. On a

$$u' = u_x x' + u_y y'$$

On peut voir la partie « champ vectoriel » (ou « champ de tangentes ») d'un problème de Cauchy comme $y' = f(x(t), y(t))$ avec $x(t) = t$. On a en particulier $f' = f_x$ et dans ce cas

$$y'' = f_x + f_y y' \quad (*)$$

en vertu du théorème ci-dessus. On a

$$\begin{aligned} y''' &= (y'')' \\ &= (f_x + f_y y')' & (*) \\ &= f'_x + (f_y y')' \\ &= f'_{xx} + (f'_{xy} y' + f_y y'') & \text{le théorème pour } f'_x \text{ et } f'_y \\ &= f_{xx} + f_{xy} y' + ((f_{yx} + f_{yy} y') y' + f_y y'') & f_{xy} = f_{yx} \text{ car } f \in C^\infty, \text{ par hypothèse} \\ &= f_{xx} + 2f_{xy} y' + f_{yy} (y')^2 + f_y y'' \\ &= f_{xx} + 2f_{xy} y' + f_{yy} (y')^2 + f_y (f_x + f_y y') & (*) \\ &= f_{xx} + 2f_{xy} y' + f_{yy} (y')^2 + f_x f_y + f_y^2 y' \end{aligned}$$

On peut maintenant calculer $y^{(iv)} = (y''')'$ en termes de dérivées partielles de f et de y' en appliquant « le théorème » et en injectant l'équation (*). Le processus se laisse itérer, on peut en particulier exprimer les dérivées d'ordre supérieur $y^{(n)} = (y^{(n-1)})'$ en termes de dérivées partielles de f et de y' , ceci pour tout n . Les dérivées partielles de f sont connues (car f est connue, partie de la donnée du problème de Cauchy), on peut donc les évaluer en (t_0, y_0) . De plus on connaît aussi la valeur $y'(t_0)$ (car c'est la condition initiale, partie de la donnée du problème de Cauchy), ce qui implique qu'on connaît $y^{(n)}(t_0)$ pour tout n sans forcément connaître $y^{(n)}$. Résultat des courses: le problème de Cauchy contient suffisamment d'information pour calculer l'expansion de Taylor d'une solution en tout degré, ceci *sans la connaître explicitement*.

Exemple. Soit $f(x, y) = x + y^2$. Approximons la solution de

$$\begin{aligned} y' &= f(t, y) \\ &= t + y^2 \\ y(0) &= 1 \end{aligned}$$

au degré 3. Supposons que $y(x)$ est une solution. L'expansion de Taylor en question est

$$\begin{aligned} y(t) &= y(0) + y'(0)(t-0) + \frac{1}{2}y''(0)(t-0)^2 + \frac{1}{3!}y'''(0)(t-0)^3 + R \\ &= y(0) + y'(0)t + \frac{1}{2}y''(0)t^2 + \frac{1}{6}y'''(0)t^3 + R \end{aligned}$$

Il nous faut donc $y(0)$, $y'(0)$, $y''(0)$ et $y'''(0)$:

$$\begin{aligned} y(0) &= 1 \\ y'(0) &= 0 + (y(0))^2 \\ &= 1 \\ y''(0) &= (f_x + f_y y')(0) \\ &= (1 + 2yy')(0) \\ &= 1 + 2y(0)y'(0) \\ &= 1 + 2 \\ &= 3 \\ y'''(0) &= (f_{xx} + 2f_{xy}y' + f_{yy}(y')^2 + f_x f_y + f_y^2 y')(0) \\ &= (0 + 0 + 2(y')^2 + 2y + (2y)^2 y')(0) \\ &= 2y'(0)^2 + 2y(0) + (2y(0))^2 y'(0) \\ &= 2 + 2 + 4 \\ &= 8 \end{aligned}$$

Du coup

$$y(t) = 1 + t + \frac{3}{2}t^2 + \frac{4}{3}t^3 + R$$

Bien sûr, à la main ça serait un peu de travail si on voulait aller au degré 30... Sur machine c'est un petit challenge sympa de calcul mathématique, car il nous faut alors concevoir un *algorithme symbolique* automatisant le processus ci-dessus. Or SAGE ne sait pas *par défaut*

- i. injecter l'équation (*);
- ii. évaluer *symboliquement* $y^{(n)}(t_0)$

(nous humains le faisons par *déduction*, ce n'est pas un processus *mécanique*). Il nous faut donc effectuer *explicitement* ces opérations à chaque itération, au moyen de *substitutions symboliques*. L'algorithme en soi n'est pas très long car SAGE sait par défaut appliquer la règle de la dérivée

composée à 2 variables (alors qu'à la main c'est assez laborieux):

Constants $t_0 \in \mathbb{R}, y_0 \in \mathbb{R}, n \in \mathbb{N}, F: \mathbb{R}^2 \rightarrow \mathbb{R}$

Symbols x, t

Functional symbols $u(t), y(t)$

$y_1 := F(t, y(t))$

$T := y_0 + (y_1[y(t) \leftarrow y_0][t \leftarrow t_0]) x$

$u(t) := y_1$

for i **in** $\text{range}(2, n+1)$:

$u := \frac{du(t)}{dt} \left[\frac{dy(t)}{dt} \leftarrow y_1 \right]$

$T := T + \frac{u[y(t) \leftarrow y_0][t \leftarrow t_0]}{i!} x^i$

return T

T est le polynôme de Taylor qu'on calcule. À l'initialisation on a

$$T = y_0 + (y_1[y(t) \leftarrow y_0][t \leftarrow t_0]) x$$

c'est à dire T à l'ordre 1. À chaque itération u est la dérivée de y à l'ordre précédent. Comme mentionné, SAGE calcule $\frac{du(t)}{dt}$ correctement par défaut. La substitution $\left[\frac{dy(t)}{dt} \leftarrow y_1 \right]$ injecte l'équation (*) dans $\frac{du(t)}{dt}$ alors que la substitution $[y(t) \leftarrow y_0][t \leftarrow t_0]$ évalue (symboliquement) $y^{(n)}(t_0)$.