ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES

Chris Peters

Université de Grenoble I
Saint-Martin d'Hères, France

20 Mai, 1996

Sommaire

Chapitre 1. Introduction

Il s'agit de rappeler quelques méthodes explicites pour résoudre des équations différentielles
On considère une équation différentielle

$$\text{(E)} \qquad \frac{dy}{dx} = f(x, y)$$

où f est une fonction continue sur un ouvert $U \subset \mathbb{R}^2$. On dit qu'une fonction dérivable $y = y(x)$ est une SOLUTION de (E) sur l'intervalle $I =]a, b[$ si $y'(x) = f(x, y(x))$ pour $x \in I$. En général on a PLUSIEURS solutions (Regarder le cas $y'(x) = f(x)$, la solution générale est une PRIMITIVE de $f(x)$, qui est bien-définie à une constante d'intégration près). Pour restreindre le nombre des solutions on étudiera plutôt le PROBLÈME DE CAUCHY (résoudre une équation différentielle avec CONDITION INITIALE):

$$\text{(C)} \qquad \begin{aligned} \frac{dy}{dx} &= f(x, y) \\ y(x_0) &= y_0 \end{aligned}$$

et on verra qu'il y a une UNIQUE SOLUTION si par exemple f est continûment dérivable en y.

§ 1. Séparation des variables

Ici $f(x, y) = a(x)/b(y)$. Une solution implicite est donnée par $\int b(y)dy = \int a(x)dx$.

1.1. Exemple. $f(x, y) = -2x/y$. On trouve $\int ydy = -2\int xdx$ et donc $\frac{1}{2}y^2 = -x^2 + C$ où C est une constante d'intégration : les solutions forment une famille d'ellipses $2x^2 + y^2 = C'$. Le problème de Cauchy admet donc une solution unique avec $I = \mathbb{R}$ si on réécrit l'équation comme $xy' = -2yx$.

1.2. Exemple. $f(x, y) = 2|y|^{\frac{1}{2}}$ avec condition initiale $y(0) = 0$. Ici il y a une infinité de solutions avec $x \geq 0$. On prend $y(x) = 0$ sur l'intervalle $[0, a]$ et ensuite $y(x) = (x - a)^2$ pour $x > a$. On peut quand même prouver que cela donne toutes les solutions.

1.3. Exemple. Si $f(x, y) = f(ax + by)$ on trouvera une solution en substituant $z = ax + by$. Par exemple, si $f = \sin^2(x - y)$ on trouve $z' = 1 - y' = 1 - \sin^2(x - y) = \cos^2 z$ et donc, si $\cos z \neq 0$ on trouve la solution $\tan(z) = x + C$, où C est une constante d'intégration et donc $y(x) = x - \arctan(x + C)$. Si $\cos(z_0) = 0$, on trouve la solution $z = z_0$ et donc $y = x + \pi/2 + k\pi$, $k \in \mathbb{Z}$.

§ 2. Équations homogènes

C'est le cas où $f = f(y/x)$. On pose $y = xz$. De là $y' = z + xz' = f(z)$ et donc $z' = \dfrac{f(z) - z}{x}$, une équation à variables séparées. Si a est une racine de $f(z) - z = 0$ on a les solutions $z = a$

et donc $y = ax$, des droites passant par 0. Si $f(z) \neq z$ avec disons $z \in]a, b[$, a, b deux racines de $f = z$ et avec $G(z)$ une primitive sur $]a, b)$ de $z \mapsto 1/(f - z)$, on trouve les solutions sous forme de famille de courbes $G(y/x) = \log(Cx)$, avec C une constante d'intégration. Ces solutions sont valables dans le secteur angulaire $a < \dfrac{y}{x} < b$, $Cx > 0$.

2.1. Exemple. $f = 1 + y/x$. Ici $1/(f - z) = 1$ et donc $G(z) = z$ et les solutions sont $y = x \log(Cx)$, valable pour $Cx > 0$.

2.2. Exemple. $f = \dfrac{y^2}{x(2y - x)} = \dfrac{z^2}{2z - 1}$. On trouve $1/(f(z) - z) = -\dfrac{2z - 1}{z(1 - z)} = -\dfrac{1}{z - 1} - \dfrac{1}{z}$ et donc $G(z) = -\log|z(1 - z)|$. Cela donne $\log|z(z - 1)| = \log C|x|^{-1}$ ou bien $y(x - y) = C'x$, une famille de coniques. Puisque l'équation s'écrit $(y - C')(y - x + C') = -C'^2$, ce sont des hyperboles passant par $(0, 0)$ avec les asymptotes $y = C'$ et $y = x - C'$. On a aussi les solutions singulières $z = 0, z = 1$ qui correspondent à $y = 0$ et $y = x$.

§ 3. Équations linéaires du premier ordre

Ce sont les équations de la forme

$$y' - a(x)y = b(x) \tag{E2}$$

avec a, b deux fonctions continues sur un intervalle I. Associée à une telle équation on a l'équation HOMOGÈNE (sans second membre) $y' - a(x)y = 0$. On a :

1 Les solutions de (E2) sont tous de la forme *solution particulière+solution générale de l'équation homogène associée*. En effet, si $y_1(x), y_2(x)$ sont deux solutions de (E2), alors leurs différence satisfait à $(y_1 - y_2)' - a(x)(y_1 - y_2) = (y_1)' - a(x)y_1 + (y_2)' - a(x)y_2 = 0$ et inversement, si y est une solution de (E2), lorsque $y_0' - a(x)y_0 = 0$ on a $(y + y_0)' - a(x)(y + y_0) = b(x)$.

2 Les solutions de l'équation homogène $y' - a(x)y = 0$ sont $e^{A(x)}$ où $A(x)$ est une primitive de $a(x)$.

3 Pour trouver une solution particulière on utilise le méthode des variation des constantes. On écrit donc $y = g(x)e^{A(x)}$ et on substitue pour trouver la solution $g(x) = \int_a^x b(t)e^{-A(t)}dt$.

3.1. Exemple. $\sin(x)y' + \cos(x)y = 2x$ donne $(\sin x \cdot y)' = 2x$ et donc $y = \dfrac{x^2 + C}{\sin x}$ pour $\sin x \neq 0$.

3.2. Exemple. ÉQUATIONS de BERNOUILLI

Ce sont les équations avec $f = p(x)y + q(x)y^m, m \neq 1$. On les ramène aux équations linéaires en multipliant par y^{-m} : $y^m y' = p(x)y^{1-m} + q(x)$ et en substituant $z = y^{1-m}$: $(1/1 - m)z' = p(x)z + q(x)$.

§ 4. Équations exactes

Ce sont des équations de la forme $a(x, y) + b(x, y)y' = 0$ avec $\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}$. On suppose que a et b soit définies sur un ouvert convexe. Dans ce cas c'est bien connu qu'il y a une fonction F tel que $\nabla(F) = (F_x, F_y)$ (le gradient de F) équivaut (a, b) : $\frac{\partial F}{\partial x} = a$, $\frac{\partial F}{\partial y} = b$. On dit que F est un potentiel. En écrivant l'équation sous la forme $adx + bdy = dF = 0$ où $dF = \frac{\partial F}{\partial x} \cdot dx + \frac{\partial F}{\partial y} \cdot dy$ on trouve une famille de solutions $F(x, y) = C$. Même si $\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}$ on peut parfois trouver un FACTEUR INTÉGRANT, c.à.d. une fonction $h(x, y)$ t.q $\frac{\partial F}{\partial x} = h \cdot a$, $\frac{\partial F}{\partial y} = h \cdot b$.

4.1. Exemple. $2x \sin y dx + (x^2 \cos y + y^2) dy = 0$ donne $F = x^2 \sin y + \frac{1}{3} y^3$.

4.2. Exemple. $y dx + x(2xy - 1) dy = 0$. On trouve un facteur intégrant h avec $\frac{dh}{dy} = 0$ en substituant : $h = -x^{-2}$ et on trouve les solutions $y^2 x - y + Cx = 0$.

§ 5. Équations linéaires

Ce sont les équations de la forme

(E). $$a_n \frac{d^n y}{dx^n} + a_{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \cdots a_0 y = b$$

Si on met $b = 0$ on obtient l'équation homogène associée (E2) et on a ici aussi :

Les solutions de (E) sont tous de la forme *solution particulière+solution générale de l'équation homogène associée* (E2). On montrera dans le chapitre 5 que les solutions de de (E2) forment un espace vectoriel de dimension n. En général il n'est pas facile de trouver des solutions. On traitera seulement le cas où les coefficients sont constants. Par substitution on vérifie que $y(x) = e^{\lambda x}$ est une solution où λ est une racine de l'équation CARACTÉRISTIQUE :

$$P(\lambda) := a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots a_0 = 0.$$

et donc, si $P(\lambda)$ a n racines réelles distinctes, alors on obtient une base pour l'espace des solutions.

Regardons en détail ce qui se passe pour $n = 2$. Donc, si $P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$ est une factorisation en facteurs linéaires il y a trois cas :

1 λ_1, λ_2 sont différents et réels. Alors $\{e^{\lambda_1 x}, e^{\lambda_2 x}\}$ est une base de solutions.

2 λ_1, λ_2 sont deux nombres complexes conjugués disons $\lambda_1 = \alpha + \mathrm{i}\beta$. Dans ce cas on a une solution complexe $e^{\lambda x} := e^{\alpha x} e^{\mathrm{i}\beta x}$ et donc deux solutions indépendantes $e^{\alpha x} \cos \beta x$ et $e^{\alpha x} \sin \beta x$.

3 $\lambda_1 = \lambda_2 =: \lambda$. On vérifie que $e^{\lambda x}$ et $x e^{\lambda x}$ sont des solutions.

Chapitre 2. Problème de Cauchy

Dans ce chapitre nous montrons l'existence et l'unicité des solutions du problème de Cauchy, localement ainsi que globalement. Bien sûr il faut faire quelques hypothèses supplémentaires. Les énoncés précis se trouvent dans le §1. Dans le §2 on montre l'existence et l'unicité locale sous l'hypothèse que $f(t, y)$ est continue et localement lipschitzienne en y. Une autre démonstration qui utilise la méthode d'Euler (cruciale dans le chapitre suivant) sera donné en §3 sous une hypothèse encore plus restrictive : f doit être continûment dérivable en t et y. Dans le §5 on montrera comment enlever cette hypothèse pour arriver à l'existence locale si f est seulement continue. Dans le §4 on montrera les résultats globaux.

§ 1. Les énoncés des résultats

Dès maintenant j'utilise plutôt les variables (t, y) dans un ouvert U du plan $\mathbb{R}^2$. On pourra regarder t comme le 'temps' et une solution comme un chemin $y(t)$ $t \in\,]a, b[= I$ dans $\mathbb{R}$ parcouru entre $t = a$ et $t = b$. Le graphe $(t, y(t))$ doit donc rester dans U et cela nous mène à la notion d'un RECTANGLE DE SÉCURITÉ. Soit $t_0 \in I$ et $y_0 \in \mathbb{R}$ tel que le point $p_0 = (t_0, y_0)$ est dans U. Soit $R \subset U$ un rectangle fermé contenu dans U et centré en p_0, disons de demi-longueur T et de demi-hauteur h. Regardons les solutions du problème de Cauchy dans R avec condition initiale $y(t_0) = y_0$. Pour simplifier nous supposons que $t_0 = 0$ et $y_0 = 0$ et donc R est un rectangle centré à l'origine. Pour une solution, il faut que $|y(t)| \leq h$ quel que soit $t \in\,] - T, T[$ (le graphe déterminé par la solution ne peut pas s'échapper de R avant $t = T$). Si c'est le cas on dit que R est un rectangle de sécurité.

Pour trouver une condition suffisante, nous regardons un rectangle R_0 fermé contenu dans U avec même centre et hauteur mais de demi-longueur T_0 arbitraire. Puisque R_0 est fermé et borné, le théorème de Heine-Borel-Lebesgue nous garantit l'existence d'un maximum fini :

$$M_0 := \max\{|f(p)|; p \in R_0\}$$

1.1. Lemme. *Si $T \leq \min\{T_0, h/M_0\}$, le rectangle*

$$R = [t_0 - T, t_0 + T] \times [y_0 - h, y_0 + h]$$

est un rectangle de sécurité.

Démonstration. Pour simplifier les notations, on suppose que $t_0 = 0$ et $y_0 = 0$. Soit τ le premier instant où $(\tau, y(\tau))$ s'échappe :

$$\tau = \inf\{t \in I; |y(t)| > h\}.$$

Par continuité on a $|y(\tau)| = h$ et donc $h = |\int_0^\tau y'(t)dt|$ = longueur total du chemin parcouru entre $t = 0$ et $t = \tau$, c.à.d. $h = |\int_0^\tau f(t, y(t))dt| \leq M_0 \cdot \tau$ et donc $h/M_0 \leq \tau$. Si $T \leq h/M_0$, le graphe ne peut donc pas s'échapper du rectangle. ■

Donc soit $T = T_0$ et $R_0 = R$ est de sécurité, soit $T = h/M_0$. Dans ce cas $T \leq h/M$, où $M = \max\{|f(p)|; p \in R\}$. Dans les deux cas R est un rectangle de sécurité au sens strict :

1.2. Définition. Soit $f : U \to \mathbb{R}$ une fonction continue sur un ouvert $U \subset \mathbb{R}^2$, $(t_0, y_0) \in U$. Un rectangle $R = [t_0 - T, t_0 + T] \times [y_0 - h, y_0 + h]$ centré en (t_0, y_0) avec $T \leq h/M$, $M = \max |f(p)|$, $p \in R$ est appelé un rectangle de sécurité au sens strict.

Utilisant la notion de rectangle de sécurité on peut énoncer les principaux résultats (EXISTENCE et UNICITÉ LOCALE) :

1.3. Théorème. (**Cauchy-Peano-Arzela**) *Si $f \in U$ est continue, le problème de Cauchy admet une solution dans un intervalle $[t_0 - T, t_0 + T]$. Plus précisément, une solution existe si T est choisi tel que $R = [t_0 - T, t_0 + T] \times [h - y_0, h + y_0] \subset U$ est un rectangle de sécurité au sens strict pour f. Si de plus f est localement lipschitzienne par rapport à y (par exemple si f est continûment dérivable par rapport à y) cette solution est unique.*

Rappel : si U est un ouvert dans $\mathbb{R}^2$ avec norme euclidienne $||..||$, on dit que $f : U \to \mathbb{R}$ est lipschitzienne avec constante de Lipschitz k si $|f(P) - f(Q)| \leq k||x - y||$ quels que soient $P, Q \in U$. On dit que f est localement lipschitzienne par rapport à y, si pour chaque $p \in U$ on peut trouver un rectangle ouvert $Q \subset U$ centré en p tel que, en fixant $t = t'$, la fonction $y \mapsto f(t, y)$ est lipschitzienne sur $Q \cap \{t = t'\}$. Si $R \subset U$ est un rectangle fermé, on peut couvrir R par un nombre fini de tels rectangles, car R est compact. Donc :

1.4. Lemme. *Si $f : U \to \mathbb{R}$ est localement lipschitzienne par rapport à y et si $R \subset U$ est un rectangle fermé, il y a un constante $k = k_R$ tel que $|f(t, y) - f(t', y)| \leq k|y - y'|$ sur R.*

Pour étudier le problème globalement sur U on a besoin de quelques définitions :

1.5. Définition.

1 Soient I_1, resp. I_2 deux intervalles de $\mathbb{R}$ sur lesquels on a les solutions f_1, resp. f_2 du problème de Cauchy. Si $I_1 \subset I_2$ et $f_2|I_1 = f_1$ on dit que f_2 est un prolongement de f_1. Si de plus $I_2 \neq I_2$ on dit que c'est un prolongement strict.

2 Une solution est dite maximale si elle n'a pas de prolongement strict.

3 Si $U = J \times \mathbb{R}$, $J \subset \mathbb{R}$ intervalle ouvert, une solution $y : J \to \mathbb{R}$ est appelée une solution globale.

1.6. Théorème.

1 Chaque solution du problème de Cauchy se prolonge en une solution maximale définie sur un intervalle ouvert.

2 Si $U = J \times \mathbb{R}$, $J \subset \mathbb{R}$ intervalle ouvert, et si pour $t \in J$ quelconque la fonction $y \mapsto f(t, y)$ est lipschitzienne de rapport $k(t)$, continue en t, alors toute solution maximale est globale.

En particulier, dans la situation du précédent théorème, il y a une unique solution globale pour le problème de Cauchy.

§ 2. Méthode de Picard

Le but est de montrer l'existence et l'unicité locale sous l'hypothèse que f est localement lipschitzienne en y. Dans le §5 on revient au cas où f est seulement continue.

L'observation fondamentale est que l'équation

$$
\text{(C)} \qquad \begin{aligned} \frac{dy}{dx} =& f(x,y) \\ y(x_0) =& y_0 \end{aligned}
$$

est équivalente à

$$
\text{(I)} \qquad y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))ds.
$$

On ne regarde que des solutions avec $t \geq t_0$. On fixe un rectangle de sécurité R au sens strict, centré en (t_0, y_0) de demi-longueur T et demi-hauteur h. Pour un tel rectangle on a :

$$
\begin{aligned} T \leq& h/M \\ \text{avec } M =& \max_R |f(x,y)|. \end{aligned}
$$

On introduit l'espace vectoriel

$$
\begin{aligned} E \ :=& \mathcal{C}^0(I, \mathbb{R}) \\ & \text{muni de la norme } ||..|| = \sup_I |..| \end{aligned}
$$

L'espace E est un espace métrique complet car une suite de Cauchy de E est une suite de fonctions continues sur I qui converge uniformément vers une fonction sur I qui est donc continue. On pose

$$
F \ := \{y \in E; ||y - y_0|| \leq h\}.
$$

C'est un sous-ensemble fermé (si $y_\nu \in F$ alors $||y_\nu - y_0|| \leq h$ ce qui implique $|| \lim_\nu y_\nu - y_0|| \leq h$) et donc aussi complet.

L'équation (I) suggère de considérer l'application

$$
\begin{aligned} \Phi \ :& F \longrightarrow E \\ & y \longmapsto y^* \ := y_0 + \int_{t_0}^{t} f(s, y(s))ds. \end{aligned}
$$

On note que $T \leq h/M$ implique que $||y^* - y_0|| \leq MT \leq h$ et donc $y^* \in F$: Φ applique F en lui-même.

Une solution de (I) est un point fixe de Φ et il est donc naturel d'appliquer le théorème suivant dont on admet la démonstration :

2.1. Théorème. *(THÉORÈME DU POINT FIXE) Soit (F, d) un espace métrique complet et $T \ : F \to F$ une application contractante : $\exists k < 1$ t.q. $d(Tx, Ty) < kd(x, y)$ quel que soient $x, y \in F$ (c.à.d. T est lipschitzienne de rapport $k < 1$). Alors T admet un point fixe et c'est le seul point fixe de T.*

2.2. Corollaire. *Si $T = \Phi^k$ (k entier et positif) est contractante, alors Φ admet un seul point fixe.*

Démonstration. (du corollaire) Soit y le point fixe de T. On considère $T(\Phi(y)) = \Phi^{k+1}y = \Phi T(y) = \Phi(y)$ et donc par unicité $y = \Phi y$. ■

Il suffit donc de montrer que Φ^k est contractante pour k suffisamment grand. Soient $y_1, y_2 \in F$ quelconques. Par le lemme 2.1.1 on a :

$$|f(s, y_1) - f(s, y_2)| \leq L \cdot |y_1 - y_2|$$

et donc (avec $y^* = \Phi(y)$ quel que soit $y \in F$) :

$$|y_1^*(t) - y_2^*(t)| \leq L \cdot (t - t_0)||y_1 - y_2||.$$

Si on applique cela pour y_1^* et y_2^* on trouve

$$\begin{aligned}
|y_1^{**}(t) - y_2^{**}(t)| &\leq L \int_{t_0}^{t} |y_1^* - y_2^*| ds \\
&\leq L \cdot \left(\int_{T_0}^{t} (s - t_0) ds \right) ||y_1^* - y_2^*|| \\
&\leq \frac{1}{2} L^2 \cdot (t - t_0)^2 ||y_1 - y_2|| \\
&\leq \frac{(LT)^2}{2} ||y_1 - y_2||.
\end{aligned}$$

Par récurrence on trouve

$$||\Phi^k y_1 - \Phi^k y_2|| \leq \frac{(LT)^k}{k!} ||y_1 - y_2||.$$

Puisque $\lim_{k\to\infty} \dfrac{a^k}{k!} = 0$ quel que soit $a \in \mathbb{R}$, l'application Φ^k est contractante pour k suffisamment grand.

§ 3. Méthode d'Euler

On subdivise l'intervalle $[t_0, T]$: $t_0 < t_1 < \ldots < t_N = T$ et on pose

$$\begin{aligned}
h_j &= t_{j+1} - t_j, \quad j = 0, \ldots, N-1 \\
h_{\max} &= \max_j h_j.
\end{aligned}$$

Le méthode d'Euler consiste à regarder la fonction linéaire par morceaux $\hat{y}(t)$ définie par récurrence par :

$$\begin{aligned}
\hat{y}(t_0) &= y_0 \\
\hat{y}_j &= \hat{y}(t_j) \\
\hat{y}(t)|[t_j, t_{j+1}] &= \hat{y}_j + f(t_j, \hat{y}_j) \cdot (t - t_j).
\end{aligned}$$

Donc $\hat{y}$ est linéaire sur l'intervalle $[t_0, t_1]$ de même pente que la vraie solution en t_0. Cela détermine y_1. Sur l'intervalle suivant $\hat{y}$ est linéaire de même pente que la vraie solution en t_1 et cela détermine y_2, etc. L'idée est que $\hat{y}$ approche la vraie solution si $h_{\max}$ tend vers zéro.

3.1. Définition. $\hat{y}$ est ϵ-approchée si

1) Le graphe de $\hat{y}$ reste dans le rectangle $R \subset U$ de sécurité,
2) Sur chaque intervalle ouvert $]t_j, t_{j+1}[$ on a $|\hat{y}' - f(t, \hat{y})| \leq \epsilon$.

On fait l'hypothèse SUPPLÉMENTAIRE :

$f \in \mathcal{C}^1(U)$, $U \subset \mathbb{R}^2$ l'ouvert de définition. On pose (R est rectangle de sécurité) :

$$M = \max_R |f(t,x)|$$
$$K = \max_R |\frac{\partial f}{\partial t}|$$
$$L = \max_R |\frac{\partial f}{\partial y}|$$

et comme d'habitude on suppose

$$\text{(sec)} \qquad T \leq \frac{h}{M}.$$

3.2. Proposition. *$\hat{y}$ est ϵ-approchée avec $\epsilon = (K + ML)h_{\max}$.*

Démonstration.

1) Par construction on a que sur l'intervalle $[t_j, t_{j+1}]$, $|\hat{y}(t) - \hat{y}_j| = (t - t_j) \cdot |f(t_j, \hat{y}_j)|$ et donc, par récurrence, si $(t_j, y_j) \in R$, le point $(t_j, \hat{y}_j)$ reste dans R et donc $|f(t_j, \hat{y}_j)| \leq M$ et $|y(t) - \hat{y}_j| \leq (t - t_j) \cdot M$. Par conséquent $|\hat{y}(t) - y_0| \leq |\hat{y}(t) - \hat{y}_j| + |\hat{y}_j - \hat{y}_{j-1}| + \cdots + |\hat{y}_1 - y_0| \leq (t - t_j) \cdot M + (t_j - t_{j-1}) \cdot M + \cdots + (t_1 - t_0) \cdot M = (t - t_0) \cdot M \leq TM \leq h$ par l'hypothèse (sec) faite sur notre rectangle de sécurité. Il suit que $(t, \hat{y}(t))$ reste dans R pour $t \leq t_{j+1}$.
2) Sur l'intervalle $]t_j, t_{j+1}[$ on a $\hat{y}' = f(t_j, y_j)$ et donc

$$\begin{aligned}
|\hat{y}_j - f(t, \hat{y}(t))| =& |f(t, \hat{y}(t)) - f(t_j, \hat{y}_j)| \\
\leq& |f(t, \hat{y}(t)) - f(t_j, \hat{y}(t))| + |f(t_j, \hat{y}(t)) - f(t_j, \hat{y}_j)| \\
\leq& K \cdot |t - t_j| + L \cdot |\hat{y}(t) - \hat{y}_j| = K \cdot |t - t_j| + L \cdot |t - t_j||f(t_j, \hat{y}_j)| \\
\leq& (K + LM)h_{\max}.
\end{aligned}$$

■

Maintenant on va comparer deux solutions approchées, l'une, $\hat{y}$ qui est ϵ-approchée et avec condition initiale $\hat{y}_0 = y_0$, l'autre $\hat{z}$ qui est ϵ'-approchée avec valeur initiale $\hat{z}(0) = z_0$. Pour faciliter l'écriture on omettra les "chapeaux" dans les notations. On pose pour $t \in I, t \neq t_1, \ldots, t \neq t_{N-1}$:

$$\begin{aligned}
\alpha(t) =& y'(t) - f(t, y(t)) \\
\beta(t) =& z'(t) - f(t, z(t)). \\
\delta(t) =& |y(t) - z(t)|.
\end{aligned}$$

On a

$$\begin{aligned}
(y(t) - y_0) - (z(t) - z_0) = (z_0 - y_0) + \int_{t_0}^{t} (\alpha(s) - \beta(s))ds \\
- \int_{t_0}^{t} ((f(s, y(s) - f(s, z(s))ds
\end{aligned}$$

et les inégalités

$$|\alpha(t)| \leq \epsilon, \quad |\beta(t)| \leq \epsilon'$$

impliquent :

$$\delta(t) \leq |y_0 - z_0| + (\epsilon + \epsilon')(t - t_0) + L \int_{t_0}^{t} \delta(s)ds.$$

Il s'agit de résoudre une telle inégalité implicite.

3.3. Lemme. *(Lemme de Gronwall) Soit* $\delta\ : [0,T] \to \mathbb{R}$ *une fonction* $\mathcal{C}^1$ *par morceaux telle que* $\exists a, b, c \in \mathbb{R}$, $c > 0$ *avec*

$$\forall t \geq 0, \quad \delta(t) \leq a + bt + c\int_0^t \delta(s)ds$$

alors

$$\forall t \geq 0, \quad \delta(t) \leq ae^{ct} + \left(e^{ct} - 1\right)\frac{b}{c}.$$

Démonstration. On soustrait $c\int_0^t \delta(s)ds$ et on multiplie par e^{-ct} :

$$\begin{aligned}\left(\delta(t) - c\int_0^t \delta(s)ds\right)e^{-ct} &= \\ \frac{d}{dt}\left[\left(\int_0^t \delta(s)ds\right)e^{-ct}\right] &\leq (a+bt)e^{-ct} \\ &= \frac{d}{dt}\left[\frac{a}{c} - \frac{(a+bt)}{c}e^{-ct} + \frac{b}{c^2}(1-e^{-ct})\right]\end{aligned}$$

et on intègre et après on multiplie par ce^{ct} :

$$c\int_0^t \delta(s)ds \leq ae^{ct} - (a+bt) + \frac{b}{c}(e^{ct}-1).$$

En somme on trouve bien :

$$\delta(t) \leq a + bt + c\int_0^t \delta(s)ds \leq a + bt + ae^{ct} - (a+bt) + \frac{b}{c}(e^{ct}-1) = ae^{ct} + \frac{b}{c}(e^{ct}-1).$$

∎

3.4. Corollaire. *On a*

$$|y(t) - z(t)| \leq |y(0) - z(0)|e^{L(t-t_0)} + \frac{\epsilon + \epsilon'}{L}(e^{L(t-t_0)} - 1).$$

Comme une première application, supposons que $\{y_\nu\}, \nu = 1, \ldots$ est une suite de fonctions ϵ_ν-approchées avec même condition initiale. Dans ce cas on trouve

$$||y_\nu - y_\mu|| \leq \frac{\epsilon_\nu + \epsilon_\mu}{L}(e^{LT} - 1)$$

et donc, si $\{\epsilon_\nu\}$ est une suite ayant zéro pour limite, les fonctions $\{y_\nu\}$ forment une suite de Cauchy dans l'espace des fonction continues sur I muni de la norme sup $|..|$. La suite converge donc uniformément vers une fonction y, continue sur I. En résumant :

3.5. Corollaire. *Soit* $\{y_\nu\}_{\nu=1}^\infty$ *une suite de fonctions* ϵ_ν*-approchées avec* $y_\nu(0) = y_0$*. Si* $\lim_{\nu\to\infty}\epsilon_\nu = 0$*, alors la suite converge uniformément vers une fonction* y*, continue sur* I

Cette fonction est une solution exacte grâce à :

3.6. Lemme. *Soit* $\{\epsilon_\nu\}_{\nu=1}^\infty$ *une suite de nombres réels avec* $\lim_{\nu\to\infty}\epsilon_\nu = 0$ *et soit* $\{y_\nu\}_{\nu=1}^\infty$ *une suite de fonctions sur* I*, linéaires par morceaux, avec la condition initiale* $y_\nu(0) = y_0$ *et telle que* $\{y_\nu\}_{\nu=1}^\infty$ *converge uniformément vers une fonction continue* y*. Alors* y *est une solution exacte.*

Démonstration. On a $|y'_\nu(t) - f(t, y_\nu(t))| \leq \epsilon_\nu$ presque partout et donc par intégration

$$(*) \qquad |y_\nu(t) - y_0 - \int_{t_0}^{t} f(s, y_\nu(s))du| \leq \epsilon_\nu(t - t_0).$$

On aussi

$$|f(s, y_\nu(s)) - f(s, y(s))| \leq L \max_I |y - y_\nu|$$

(le max existe car I est compact et y est continue). Le membre droit tend vers zéro lorsque ν tend vers l'infini, car y_ν tend uniformément vers y. En prenant la limite dans (*) quand ν tend vers l'infini, on obtient que $|y(t) - y_0 - \int_{t_0}^{t} f(s, y(s))ds| = 0$ et donc $y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))ds$ est bien dérivable et est une solution du problème de Cauchy (se reporter à la formule (I) au début du §2 et la discussion après). ■

En conclusion on a :

3.7. Théorème. *Soit $\{\epsilon_\nu\}_{\nu=1}^{\infty}$ une suite avec $\lim_{\nu\to\infty} \epsilon_\nu = 0$. Soit y_ν la fonction linéaire par morceaux sur I provenant de la méthode d'Euler pour une subdivision de I telle que $h_{\max} \leq \dfrac{\epsilon_\nu}{K + LM}$. Sous l'hypothèse que $f \in \mathcal{C}^1(U)$ la limite $y = \lim_{\nu\to\infty} y_\nu$ est dérivable et résout le problème de Cauchy.*

Démonstration. Par la proposition 2, y_ν est $(K + LM)h_{\max}$-approchée. L'hypothèse implique $(K + LM)h_{\max} \leq \epsilon_\nu$ et donc y_ν est certainement ϵ_ν-approchée. Par le corollaire 5 et l'hypothèse sur les ϵ_ν, la suite de fonctions y_ν converge uniformément vers une fonction continue y. Par le lemme 6 cette fonction est une solution du problème de Cauchy. ■

3.8. Corollaire de la méthode. *La solution $y = y(t; y_0)$ est unique et dépend de façon lipschitzienne de la condition initiale y_0 :*

$$||y(-, y_0) - y(-, z_0)|| \leq |y_0 - z_0| \cdot e^{LT}.$$

En particulier, $y(-, t_0)$ dépend continûment de t_0.

Démonstration. Pour chaque solution z qui est ϵ'-approchée, l'inégalité du Corollaire 4 donne l'estimation

$$||y - z|| \leq \frac{\epsilon'}{L}(e^{LT} - 1)$$

et en prenant $\epsilon' = 0$ on a l'unicité.

Aussi, comparant deux solutions exactes, l'inégalité du Corollaire 4 dit :

$$||y - z|| \leq |y_0 - z_0| \cdot e^{LT}.$$

■

§ 4. Solutions globales

Le but est de montrer le théorème 2.1.6.

1) Tout solution du problème de Cauchy se prolonge en une solution maximale définie sur un intervalle ouvert.

Démonstration. Dans cette démonstration on pose $|ab|$ pour un intervalle entre a et b fermé, ouvert, demi-ouvert. Dans le cas ouvert à gauche $a = -\infty$ est possible, pour un intervalle ouvert à droite $b = \infty$ est possible. Il suffit de considérer les extensions à droite. Soit $I = |ab_1|$ l'intervalle de début avec solution $y = y_1$. On pose

$$c_2 = \sup\{c \in \mathbb{R}; y_1 \text{ se prolonge sur } |ac[\}.$$

Si c_2 est fini on choisit un prolongement y_2 sur l'intervalle $|a, b_2[$ avec

$$c_2 - b_2 < \frac{1}{2}.$$

Si $c_2 = \infty$ on choisit un prolongement y_2 sur $|a, b_2[$ avec

$$b_2 > 2.$$

Par récurrence on construit

$$I_1 \subset I_2 \subset \cdots I_k = |ab_k[\cdots$$

tel que $y_k \ : I_k \to \mathbb{R}$ est un prolongement de $y_{k-1} \ : I_{k-1} \to \mathbb{R}$ et avec

$$c_k = \sup\{c \in \mathbb{R}; y_{k-1} \text{ se prolonge sur } |ac[\}.$$

Si $c_k = \infty$ on peut supposer que

$$b_k > k$$

et si c_k est fini, on peut supposer

$$c_k - b_k < \frac{1}{k}.$$

Si toujours $c_k = \infty$, pour chaque b_k on a $b_k > k$ et on pose $b_\infty = \infty$. Sinon, c_k est fini pour un certain rang k_0, et reste fini pour $k \geq k_0$, car les c_k forment une suite décroissante puisque chaque prolongement de y_k est aussi un prolongement de y_{k-1}. On a à partir de k_0 :

$$b_k \leq b_{k+1} \ldots \leq c_{k+1} \leq c_k$$

et on pose

$$b_\infty = \lim_{k\to\infty} b_k = \lim_{k\to\infty} c_k.$$

Soit I l'intervalle $|a, b_\infty|$ et y le prolongement commun des solutions y_k, éventuellement prolongé à b_∞ si possible. Le prolongement est une solution maximale : si z prolonge y sur $|a, c|$, alors $\forall k$, y est un prolongement de y_k et donc $c \leq c_{k+1} \Rightarrow c \leq \lim c_k = b_\infty$ et donc $c = b_\infty$.

Finalement, l'intervalle $|ab|$ est forcément ouvert à cause du théorème d'existence des solutions locales : si l'intervalle est $|ab]$ on pourrait continuer la solution en un intervalle autour de b ce qui contredit la maximalité.

■

2) Si $U = J \times \mathbb{R}$, où J est un intervalle ouvert et $f(t, y)$ est lipschitzienne en y de rapport $k(t)$, $k \ : J \to \mathbb{R}$ continue, alors une solution maximale est globale, c.à.d. est définie sur J tout entier.

Démonstration. On reprend la démonstration du §2. Pour chaque sous-intervalle *fermé* $I \subset J$, f est k-lipschitzienne avec $k = \max_I k(t)$. Puisque $U = J \times \mathbb{R}$ on peut choisir un rectangle $I \times [-h + y_0, y_0 + h]$ de demi-hauteur quelconque et donc dans la preuve de §2 on peut utiliser l'espace E au lieu de F. La démonstration produit une unique solution exacte valable sur chaque sous-intervalle fermé de J. Il existe alors un prolongement sur J tout entier et ce prolongement est unique partout. ∎

§ 5. Compléments

Il s'agit de prouver le théorème 2.1.3 sous l'hypothèse que f est continue. On a besoin du *module de continuité* de f sur un ensemble compact $K \subset U$:

5.1. Définition. Soit $f : U \to \mathbb{R}$ continue, U ouvert de $\mathbb{R}^2$. Soit K sous-ensemble compact de U. Alors le *module de continuité* de f est

$$\omega_{f|K}(r) = \max\{|f(t,y) - f(t',y')|; (t,y),(t',y') \in K, \quad |t-t'| + |y-y'| < r\}$$

Conséquence.

$$|f(t,y) - f(t',y')| \le \omega_{f|K}(R) \quad \text{pour les points } (t,x),(t',x') \in K, \quad \textit{avec } |t-t'| + |y-y'| < R$$

L'implication découle du fait que la fonction $\omega_{f|K}(r)$ est *croissante*. Elle prend des valeurs *finies*, car la fonction continue $((t,y),(t',y')) \mapsto |f(t,y) - f(t',y')|$ admet un maximum fini sur l'ensemble compact $K \times K$. De plus, la fonction $f|K$ étant uniformément compacte, $\forall \epsilon > 0$, $\exists r > 0$ tel que $|f(t,y) - f(t',y')| < \epsilon$ quels que soient les points $(t,y),(t',y')$ avec $|t-t'| + |y-y'| < r$. Autrement dit :

$$\text{(MC)} \qquad \lim_{r \to 0} \omega_{f|K}(r) = 0.$$

On reprend la démonstration donnée au paragraphe 3. On remplace la Prop 2. par :

$\hat{y}$ est ϵ-approchée avec $\epsilon = \omega_{f|R}((M+1)h_{\max})$.

Cela découle immédiatement du fait que

$$\begin{aligned} |\hat{y}(t) - \hat{y}(t_j)| =& |t - t_j||f(t_j, y_j)| \le |t - t_j| \cdot M \le h_{\max} \cdot M \\ &\Rightarrow \forall (t,x),(t',x') \in R \quad |t-t'| + |y-y'| \le h_{\max} \cdot (M+1) \end{aligned}$$

et donc

$$|\hat{y}_j - f(t, \hat{y}(t))| = |f(t, \hat{y}(t)) - f(t_j, \hat{y}_j)| \le \omega_{f|R}(h_{\max} \cdot (M+1)).$$

Il s'ensuit que si on choisit une suite de subdivisions de I telle que $h^{(\nu)}_{\max} \to 0$ quand $\nu \to \infty$, les approximations y_ν produites par le méthode d'Euler sont ϵ_ν-approchées avec $\epsilon_\nu \to 0$. On ne peut pas appliquer le lemme de Gronwall pour arriver au Corollaire 5. On doit remplacer la preuve de cette corollaire par une application du

Théorème de Ascoli. (Admis) *Soit E une espace métrique complet et $y_\nu : E \to \mathbb{R}$ une suite d'applications k-lipschitziennes où k est une constante donnée. Alors on peut extraire de y_ν une sous-suite qui converge uniformément vers une fonction k-lipschitzienne.*

Dans notre cas les fonctions y_ν provenant de la méthode d'Euler sont toutes lipschitziennes de rapport $k = M$. Éventuellement il faudra remplacer y_ν par une sous-suite pour arriver à une limite y qui est M-lipschitzienne et donc continue.

Pour terminer la démonstration il suffit de remplacer la démonstration du lemme 6 en utilisant l'estimation

$$|f(s, y_\nu(s)) - f(s, y(s))| \leq \omega_{f|K}(\delta_\nu),$$

où $\delta_\nu = \max_I |y(t) - y_\nu(t)|$.

$\blacksquare$

Chapitre 3. Méthodes numériques

On a déjà introduit une méthode numérique : la méthode d'Euler. En général, une méthode numérique est une procédure récurrente pour calculer des valeurs d'une solution approchée, prises dans les points d'une subdivision de l'intervalle de départ. Ici on veut étudier les méthodes à un pas c.à.d. pour calculer la valeur dans un point de la subdivision, la connaissance de la valeur dans le point précédent seule est nécessaire. Nous allons formuler des conditions qui garantissent que ces valeurs approchées approchent une solution exacte, lorsque la subdivision est rendue de plus et plus fine.

§ 1. Introduction

Il s'agit de trouver des valeurs approximatives y_i du problème de Cauchy dans les points t_i, $i = 0, \dots N$ d'une subdivision de l'intervalle I du départ par une méthode qui produit y_{k+1} lors de y_k (méthode à un pas).

Plus précisément, si

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

est la subdivision de $I = [t_0, t_0 + T]$, en posant $h_i = t_{i+1} - t_i$, une méthode à un pas est une fonction

$$\Phi \ : I \times J \times [0, \delta] \to \mathbb{R}, \quad J \subset \mathbb{R} \quad \text{un intervalle ouvert}, \quad \delta > 0$$

telle que

$$y_{i+1} = y_i + h_i \Phi(t_i, y_i, h_i), \quad 0 \leq i \leq N.$$

1.1. Exemples.

1. La méthode d'Euler. Ici $\Phi(t, y, h) = f(t, y)$.
2. La méthode du point milieu. L'idée est que la pente d'une solution $z(t)$ en $z(t+h/2)$ approxime bien la pente de la corde de $(t, z(t))$ à $(t+h, z(t+h))$. Il faut approximer $z(t+h/2)$ par une fonction qui est calculable lors de $z(t)$, par exemple par $z(t) + \frac{1}{2} h f(t, y)$ (l'approximation suggérée par la formule de Taylor). On arrive alors à

$$\Phi(t, y, h) = f\left(t + \frac{1}{2}h, y + \frac{1}{2}h \cdot f(t, h)\right).$$

L'algorithme s'écrit plus explicitement :

$$\begin{aligned} y_{i+\frac{1}{2}} &= y_i + \frac{h_i}{2} f(t_i, y_i) \\ p_i &= f\left(t_i + \frac{h_i}{2}, y_{i+\frac{1}{2}}\right) \\ y_{i+1} &= y_i + h_i p_i. \end{aligned}$$

Pour ce qui suit, la remarque suivante est cruciale :

1.2. Remarque. Si $z(t)$ résout le problème de Cauchy et si f est de classe $\mathcal{C}^k$ on peut calculer les dérivées $z^{(j)}(t)$ de $z(t)$ jusqu'à $j = k+1$. En fait, on pose

$$\begin{aligned} f^{[0]}(t,z) =& f(t,z) \\ f^{[k]}(t,z) =& \frac{\partial f^{[k-1]}}{\partial t}(t,z) + \frac{\partial f^{[k-1]}}{\partial y}(t,z) \cdot f(t,z),\ k = 1,2,\dots \end{aligned}$$

et on vérifie que

$$z^{(k+1)} = f^{[k]}(t, z(t)).$$

Donc la formule de Taylor pour $z(t)$ s'écrit :

$$\begin{aligned} z(t+h) &= z(t) + \sum_{k=1}^{p} \frac{1}{k!} h^k z^{(k)}(t) + O(h^{p+1}) \\ &= z(t) + \sum_{k=1}^{p} \frac{1}{k!} h^k f^{[k-1]}(t) + O(h^{p+1}) \end{aligned}$$

et on définit la méthode de Taylor d'ordre p par :

$$\Theta_p(t,y,h) = \sum_{k=1}^{p} \frac{1}{k!} h^{k-1} f^{[k-1]}(t,y)$$

et donc on obtient l'estimation cruciale :

$$z(t+h) - z(t) = h \cdot \Theta_p(t, z(t)) + O(h^{p+1}).$$

§ 2. Notions de base

2.1. Définition.

1. L'erreur de consistance e_i est

$$e_i = z(t_{i+1}) - z(t_i) - h_i \Phi(t_i, z(t_i), h_i)$$

   c.à.d. la différence entre l'approximation $y_{i+1} = y_i + h_i \Phi(t_i, y_i, t_i)$ et la vraie solution $z(t_{i+1})$ en supposant que $y_i = z(t_i)$.

2. La méthode est consistante, si pour toute solution exacte $z(t)$, $\sum_i |e_i|$ tend vers zéro quand $h_{\max}$ tend vers zéro.

3. La méthode est convergente, si pour chaque solution exacte $z(t)$ on a $\max_{0 \le i \le N} |y_i - z(t_i)| \to 0$ quand $h_{\max}$ tend vers zéro.

4. La méthode est stable avec constante de stabilité S, si pour la suite y_i^* définie par récurrence à partir de y_0^* (avec $\epsilon_i \in \mathbb{R}$, $i = 0, \dots, N$ quelconque) :

$$y_{i+1}^* = y_i^* + h_i \cdot \Phi(t_i, y_i^*, h_i) + \epsilon_i$$

on a

$$\max_i |y_i^* - y_i| \leq S \cdot |y_0^* - y_0| + \sum_i |\epsilon_i|.$$

5. La méthode Φ est d'ordre $\geq p$ si f et Φ sont de classe $\mathcal{C}^p$ et s'il y a une constante $C \geq 0$ telle que pour chaque subdivision de I l'erreur de consistance relative à z vérifie :

$$|e_i| \leq C h_i^{p+1}, \quad i = 0, \ldots, N.$$

Si la méthode n'est pas d'ordre $> p$ on dit qu'elle est d'ordre p.

2.2. Exemples.

1. La méthode d'Euler est d'ordre ≥ 1 :

$$\begin{aligned} e_i &= z(t_i + h_i) - (z(t_i) + h_i z'(t_i)) = \frac{1}{2} h_i^2 z''(t_i) + O(h_i^3) \\ &= \frac{1}{2} h_i^2 f^{[1]}(t_i, y_i) + O(h_i^3). \end{aligned}$$

En utilisant Taylor avec reste de Lagrange, on trouve :

$$e_i = \frac{1}{2} h_i^2 f^{[1]}(\tau, z(\tau)), \quad \tau \in]t_i, t_{i+1}[$$

et donc :

$$\begin{aligned} \sum_i |e_i| \leq & \frac{1}{2} \left(\sum_i h_i^2 \right) \max_I |f^{[1]}(t, z(t))| \\ \leq & \frac{1}{2} h_{\max} \left(\sum_i h_i \right) \max_I |f^{[1]}(t, z(t))| \\ = & \frac{1}{2} h_{\max} T \max_I |f^{[1]}(t, z(t))|. \end{aligned}$$

Donc la méthode est consistante. On verra plus tard qu'une méthode d'ordre ≥ 1 est consistante.

2. La méthode de Taylor est d'ordre $\geq p$:

$$\begin{aligned} e_i &= z(t_{i+1}) - y_{i+1} = z(t_i + h_i) - \sum_{k=1}^{p} f^{[k-1]}(t_i, y_i) \\ &= \frac{1}{(p+1)!} h_i^{p+1} f^{[p]}(t_i, h_i) + O(h_i^{p+2}). \end{aligned}$$

3. Pour la méthode du point milieu on calcule e_i comme la somme de deux contributions :

$$\begin{aligned} \epsilon_i &= z(t_{i+1}) - z(t_i) - h_i z'(t_i + \frac{1}{2} h_i) = \frac{1}{24} h_i^3 z'''(t_i) + O(h_i^4) \\ \epsilon_i' &= h_i z'(t_i + \frac{1}{2} h_i) - (y_{i+1} - z(t_i)) \\ &= h_i \left(f(t_i + \frac{1}{2} h_i, z(t_i + \frac{1}{2} h_i)) - f(t_i + \frac{1}{2} h_i, y_{i+\frac{1}{2}}) \right). \end{aligned}$$

Pour ϵ_i on trouve donc :

$$\epsilon_i = \frac{1}{24} h_i^3 f^{[2]}(t_i, y_i) + O(h_i^4).$$

D'autre part calculons d'abord

$$\begin{aligned} z(t_i + \frac{1}{2}h_i) - y_{i+\frac{1}{2}} &= z(t_i + \frac{1}{2}h_i) - \left(z(t_i) + \frac{1}{2}h_i z'(t_i) \right) \\ &= \frac{1}{8}h_i^2 z''(t_i) + O(h_i^3) = \frac{1}{8} f^{[1]}(t_i, y_i) + O(h_i^3) \end{aligned}$$

et utilisons ce-là pour calculer ϵ_i' avec le théorème des accroissements finis appliqué à $y \mapsto f(t_i + \frac{1}{2}h_i, y)$:

$$\begin{aligned} f(t_i + \frac{1}{2}h_i, z(t_i + \frac{1}{2}h_i)) - &f\left(t_i + \frac{1}{2}h_i, y_{i+\frac{1}{2}}) \right) \\ &= f_y'(t_i + \frac{1}{2}h_i, \eta) \left(z(t_i + \frac{1}{2}h_i) - y_{i+\frac{1}{2}} \right), \quad \eta \in]y_i, y_{i+1}[\\ &= \Big(f_y'(t_i, y_i) + O(h_i) \Big) \left(\frac{1}{8} h_i^2 f^{[1]}(t_i, y_i) + O(h_i^3) \right) \end{aligned}$$

et donc

$$\epsilon_i' = \frac{1}{8} h_i^3 f_y' f^{[1]}(t_i, y_i) + O(h_i^4).$$

Rassemblant les deux contributions nous trouvons :

$$e_i = \epsilon_i + \epsilon_i' = \frac{1}{24} h_i^3 \left(f^{[2]} + 3 f_y' f^{[1]} \right) (t_i, y_i) + O(h_i^4).$$

Il s'ensuit que le méthode est d'ordre ≥ 2.

Pour déterminer l'ordre d'une méthode Φ on la compare avec la méthode de Taylor :

$$\begin{aligned} e_i &= z(t_{i+1}) - z(t_i) - h_i \Phi(t_i, y_i, h_i) \\ z(t_{i+1}) - z(t_i) &= h_i \left(\Theta_p(t_i, y_i, h_i) + \frac{h_i^p}{(p+1)!} f^{[p+1]}(t_i, y_i) \right) + O(h_i^{p+2}) \\ h_i \Phi(t_i, y_i, h_i) &= h_i \left(\Phi(t_i, y_i, 0) + h_i \frac{\partial \Phi}{\partial h}(t_i, y_i, 0) + \cdots \frac{1}{p!} h_i^p \frac{\partial^p \Phi}{\partial h^p}(t_i, y_i, 0) \right) + O(h_i^{p+2}). \end{aligned}$$

Donc :

2.3. Lemme. *La méthode Φ est d'ordre $\geq p$ si et seulement si*

$$\frac{\partial^{(j)} \Phi}{\partial h^j}(t, y, 0) = \frac{1}{j+1} f^{[j]}(t, y), \quad j = 0, \ldots, p-1$$

et sous cette hypothèse on a

$$e_i = \frac{1}{p!} h_i^{p+1} \left(\frac{1}{p+1} f^{[p]}(t_i, y_i) - \frac{\partial^p \Phi}{\partial h^p}(t_i, y_i, 0) \right) + O(h_i^{p+2}).$$

§ 3. Théorèmes fondamentaux

3.1. Théorème. *Si Φ est d'ordre ≥ 1 elle est consistante.*

Démonstration. Par le théorème des accroissements finis on peut trouver $\tau \in]t_i, t_{i+1}[$ tel que

$$z(t_{i+1}) - z(t_i) = h_i z'(\tau) = h_i f(\tau, z(\tau))$$

et donc

$$e_i = z(t_{i+1}) - z(t_i) - h_i \Phi(t_i, z(t_i), h_i) = h_i \left(f(\tau, z(\tau)) - \Phi(t_i, z(t_i), h_i) \right).$$

La méthode étant d'ordre ≥ 1 on a

$$f(\tau, z(\tau)) = \Phi(\tau, z(\tau), 0)$$

et donc

$$e_i = h_i \left(\Phi(\tau, z(\tau), 0) - \Phi(t_i, z(t_i), h_i) \right).$$

La fonction $(t, h) \mapsto \Phi(t, z(t), h)$ est uniformément continue sur $[t_0, t_0 + T] \times [0, \delta]$ et donc :

$$\forall \epsilon > 0, \quad \exists \eta > 0, \quad h_{\max} \leq \eta \Rightarrow |\Phi(\tau, z(\tau), 0) - \Phi(t_i, z(t_i), h_i)| < \epsilon.$$

et il s'ensuit que

$$|e_i| = h_i \left| \Phi(\tau, z(\tau), 0) - \Phi(t_i, z(t_i), h_i) \right| \leq \epsilon h_i$$

et donc $\sum_i |e_i| \leq \epsilon T$ tend vers zéro quand $h_{\max}$ tend vers zéro. ∎

3.2. Théorème. *Si Φ est Λ-lipschitzienne en y elle est stable avec constante de stabilité $e^{\Lambda T}$.*

Démonstration. On pose

$$\Theta_i = |y_i^* - y_i|, \quad i = 0, \ldots, N.$$

Par hypothèse

$$\begin{aligned} \Theta_{i+1} = & |y_{i+1}^* - y_{i+1}| \leq |y_i^* - y_i| + h_i \cdot \Lambda \cdot |y_i^* - y_i| + \epsilon_i \\ & \leq (1 + \Lambda h_i)\Theta_i + |\epsilon_i| \\ & \leq e^{\Lambda h_i} + |\epsilon_i|. \end{aligned}$$

On vérifie alors par récurrence sur i que

$$\Theta_i \leq e^{\lambda(t_i - t_0)}\Theta_0 + \sum_{k=0}^{i-1} e^{\Lambda(t_i - t_{k+1})}|\epsilon_k|.$$

Cela se vérifie comme suit :

$$\begin{aligned} \Theta_{i+1} & \leq e^{\Lambda(t_{i+1} - t_i)}\Theta_i + |\epsilon_i| \\ & \leq e^{\Lambda(t_{i+1} - t_0)}\Theta_0 + \sum_{k=0}^{i-1} e^{\Lambda(t_i - t_{k+1})}|\epsilon_k| + |\epsilon_i|. \end{aligned}$$

Par conséquent :

$$\max_i |y_i^* - y_i| \leq e^{\Lambda T} \left(|y_0^* - y_0| + \sum_i |\epsilon_i| \right).$$

∎

3.3. Théorème. *Une méthode stable et consistante est convergente*

Démonstration. On prend $y_i^* = z(t_i)$ et $\epsilon_i = e_i$, l'erreur de consistance. Par stabilité

$$\max_i |z_i - z(t_i)| \leq S \cdot \left(\sum_i |e_i|\right)$$

et la consistance implique que $\max_i |z_i - z(t_i)|$ tend vers zéro quand $h_{\max}$ tend vers zéro, c.à.d. la méthode converge. ∎

3.4. Exemples.

Lorsque f est k-lipschitzienne en y, les méthodes d'Euler et du point milieu sont convergentes : pour la méthode d'Euler $\Phi(t, y, h) = f(t, y)$ est k-lipschitzienne en y et donc stable, pour la méthode du point milieu avec $\Phi(t, y, h) = f\left(t + \frac{1}{2}h, y + \frac{1}{2}hf(t, y)\right)$ on déduit que :

$$\begin{aligned}
|\Phi(t, y, h) - \Phi(t, y', h)| &\leq k\left|\left(y + \frac{1}{2}hf(t, y)\right) - \left(y' - \frac{1}{2}hf(t, y')\right)\right| \\
&\leq k\left(|y - y'| + \frac{1}{2}h|f(t, y) - f(t, y')|\right) \\
&\leq k\left(|y - y'| + \frac{1}{2}hk(y - y')\right) = k\left(1 + \frac{1}{2}hk\right)|y - y'|
\end{aligned}$$

et donc Φ est $k\left(1 + \frac{1}{2}h_{\max}k\right)$-lipschitzienne et donc stable. Les deux méthodes étant d'ordre ≥ 1 et donc consistante, il s'ensuit qu'elles sont convergentes.

§ 4. L'ordre de l'approximation

Il s'agit d'estimer les erreurs de consistance et de l'erreur globale à partir des connaissances des dérivés $f^{[p]}$ et $\frac{\partial^p \Phi}{\partial h^p}$. On reprend les calculs du Lemme 2.3 en utilisant la formule de Taylor avec reste de Lagrange :

$$\exists \tau \in]t_i, t_{i+1}[, \ \eta_i \in]0, h_i[\quad \text{t.q. } e_i = \frac{1}{p!}h_i^{p+1}\left(\frac{1}{p+1}f^{[p]}(\tau, z(\tau)) - \frac{\partial^p \Phi}{\partial h^p}(t_i, z(t_i), \eta_i)\right).$$

Si on pose :

$$C = \max_R \left|\frac{1}{(p+1)!}f^{[p]}(t, y)\right| + \max_{R \times [0, h_{\max}]} \left|\frac{1}{p!}\frac{\partial^p \Phi(t, y, h)}{\partial h^p}\right|$$

on a l'estimation

$$|e_i| \leq h_i^{p+1}C \leq h_i h_{\max}^p C$$

Pour l'erreur totale on a donc

$$\sum_i |e_i| \leq h_{\max}^p TC$$

et pour un méthode Φ d'ordre $p \geq 1$ qui est k-lipschitzienne en y, on trouve pour l'erreur d'approximation :

$$\max_i |y_i - z(t_i)| \leq e^{kT} \sum_i |e_i| \leq CTe^{kT}h_{\max}^p$$

et cette erreur est bien calculable.

4.1. Exemples.

1. Méthode d'Euler. On a
$$C = \max_R \frac{1}{2}|f^{[1]}(t,y)|$$
et donc, pour l'erreur d'approximation
$$\leq \frac{1}{2}Te^{kT}h_{\max}^2 \max_R |f^{[1]}|.$$

2. Méthode du point milieu. Par le calcul de e_i déjà fait on trouve bien une estimation
$$|e_i| \leq \frac{1}{24}h_i^3 \max_R |f^{[2]} + 3f_y' f^{[1]}|$$
et donc
$$\max_i |y_i - z(t_i)| \leq \frac{1}{24}h_{\max}^3 Te^{kT} \max_R |f^{[2]} + 3f_y' f^{[1]}|.$$

Chapitre 4. Systèmes d'équations linéaires

On va remplacer la variable y par un vecteur Y, ce qui est nécessaire pour résoudre des équations d'ordre ≥ 2. Le but principal est d'étudier les équations linéaires $Y' = AY + B$ et appliquer les théorèmes du chapitre 2.

§ 1. Généralités

Dans l'approche du problème de Cauchy on peut remplacer la variable y par un vecteur $Y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ et la valeur absolue $|..|$ de $\mathbb{R}$ par la norme euclidienne $||..||$ de $\mathbb{R}^m$. La fonction $f(t, y)$ peut être remplacée par une application continue

$$f \ : U \to \mathbb{R}^m, \quad U \subset \mathbb{R} \times \mathbb{R}^m \text{ ouvert}$$

et donc le problème de Cauchy s'écrit sous forme d'un système :

$$\text{(C)} \qquad \begin{aligned} \frac{dY}{dt} &= f(t, Y) \\ Y(t_0) &= Y_0 \end{aligned}$$

ou plus explicitement

$$\text{(Cs)} \qquad \begin{aligned} \frac{dy_1}{dt} &= f_1(t, y_1, \ldots y_m) \\ \frac{dy_2}{dt} &= f_2(t, y_1, \ldots y_m) \\ \vdots \quad & \qquad \vdots \\ \frac{dy_m}{dt} &= f_m(t, y_1, \ldots y_m) \\ y(t_0) &= y_0 \end{aligned} \; .$$

Un cas particulier : une équation d'ordre p disons

$$y^{(p)} = g(t, y, y', \ldots, y^{(p-1)})$$

(ici $g \ : U \to \mathbb{R}^m$ application continue et U ouvert de $\mathbb{R} \times \underbrace{\mathbb{R}^m \times \ldots \times \mathbb{R}^m}_{p \text{ fois}}$) mène à un système :

$$\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{p-1} \end{pmatrix}' = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ g(t, y_0, \ldots, y_{p-1} \end{pmatrix}.$$

Il faut réaliser que la théorie des chapitres 2 et 3 restent vraie pour les systèmes. Il s'agit simplement de remplacer la notion de "rectangle de sécurité" par la notion de "cylindre de sécurité", notion obtenue en remplaçant la valeur absolue de $\mathbb{R}$ par la norme euclidienne de $\mathbb{R}^m$. On va appliquer cette théorie dans le cas particulier d'un système linéaire.

On dit qu'on a *un système linéaire* si la fonction f est une application affine en Y, c.à.d., on a :

$$Y' = A(t)Y + B(t)$$

avec une fonction continue $A\ : J =]a,b[\to M_m$ (les matrices de taille m) et une fonction continue $B\ : J \to \mathbb{R}^m$. Donc ici $U = J \times \mathbb{R}^m$.

Une *équation linéaire d'ordre* p :

$$y^{(p)} = a_{p-1}(t)y^{(p-1)} + \ldots + a_0(t)y + b(t)$$

donne le système linéaire :

$$\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{p-2} \\ y_{p-1} \end{pmatrix}' = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \\ a_0 & a_1 & \cdots & a_{p-2} & a_{p-1} \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{p-2} \\ y_{p-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b \end{pmatrix}.$$

Traitons maintenant les systèmes linéaires en généralité en appliquant les théorèmes du chapitre 2. Puisque $f(t,Y) = A(t)Y + B(t)$ est affine en Y, cette fonction est lipschitzienne en Y avec constante de Lipschitz

$$k(t) = ||A(t)||_\infty = \max_{x \in S^{m-1}} ||A(x)||$$

continue en t et le théorème 2.5. dit que le problème de Cauchy admet une unique solution maximale qui est globale. On pose maintenant :

$$\mathrm{Sol} = \{Y \in \mathcal{C}^1(J, \mathbb{R}^m); Y' = A(t)Y\}.$$

Donc Sol est l'ensemble des solutions du système homogène associé. Soit

$$Y_0(t) \quad \text{solution de (C).}$$

Alors on a :

1.1. Théorème. Sol *est un espace vectoriel de dimension* m *et l'espace affine* $Y_0(t) + \mathrm{Sol}$ *consiste en les solutions du système linéaire inhomogène* $Y' = A(t)Y + B(t)$.

Démonstration. Soit $Y(t)$ la solution unique de (C) avec condition initiale $Y(t_0) = Y_0 \in \mathbb{R}^m$ donnée. Alors l'application $Y(t) \mapsto Y_0$ établit un isomorphisme de Sol à $\mathbb{R}^m$. Si $Y_0(t), Y_1(t)$ sont deux solutions de l'équation inhomogène, alors la différence est une solution de l'équation homogène associée et donc $Y_1(t) - Y_0(t) \in \mathrm{Sol}$. Aussi pour une solution $Y(t)$ quelconque du système homogène, la fonction $Y(t) + Y_0(t)$ est une solution du système inhomogène. On peut bien résumer tout cela en disant que les solutions du système homogène forment l'espace affine $Y_0(t) + \mathrm{Sol}$. ■

§ 2. Systèmes linéaires homogènes à coefficients constants

Ici on prend $J = \mathbb{R}$ et donc $U = \mathbb{R} \times \mathbb{R}^m$. Les uniques solutions du problème de Cauchy se prolongent sur $\mathbb{R}$ tout entier.

Considérons d'abord le cas d'une variable. On sait qu'une solution de $y' = ay$ et donnée par e^{at}. Donc, pour une matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & 0 & \lambda_m \end{pmatrix}$$

on trouve que la solution générale est donnée par

$$e^{Dt} Y_0, \quad e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & 0 & e^{\lambda_m t} \end{pmatrix}, \quad Y_0 \in \mathbb{R}^m.$$

Nous voulons généraliser cela en introduisant les exponentielles

$$e^A = I + A + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!} + \ldots, \quad A \in M_m.$$

On a les propriétés suivantes :

2.1. Lemme.

1) *La série de e^A converge normalement,*
2) *Pour $A = D$ diagonale la définition coïncide avec celle ci-dessus,*
3) *Si A et B commutent $e^A e^B = e^{A+B} = e^B e^A$,*
4) *Pour $P \in M_m$ inversible $e^{PAP^{-1}} = Pe^A P^{-1}$,*
5) *On a $\left(e^{At}\right)' = Ae^{At}$.*

On laisse ces propriétés aux lecteurs. On note que la propriété 4) dit qu'on peut réduire le calcul de e^A en transformant A en une matrice convenable. Par exemple, si A est diagonalisable, disons $A = PDP^{-1}$ avec D diagonale, alors $e^A = Pe^D P^{-1}$ se déduit immédiatement de la matrice diagonale e^D ci-dessus. Dans la pratique on calcule les λ_i et la matrice P en même temps en trouvant les valeurs propres et les vecteurs propres correspondants. La matrice P est la matrice ayant ces vecteurs propres comme colonnes.

On note finalement que la propriété 5) est utile pour résoudre l'équation $Y' = AY$. Une solution générale est donnée par $e^{At}Y_0$ car $(e^{At}Y_0)' = Ae^{At}Y_0$.

Le problème se pose pour le calcul de e^A si A n'est pas diagonalisable. Il faut supposer qu'on travaille avec des coefficient complexes. Dans la théorie précédente (dans le cas linéaire) on peut bien remplacer $\mathbb{R}^m$ par $\mathbb{C}^m$ et considérer $A \in M_m(\mathbb{C})$, $B \in \mathbb{C}^m$. L'avantage est qu'on peut se servir du théorème de Jordan pour trouver une base de $\mathbb{C}^m$ telle que la matrice obtenue de A après changement de base se décompose en blocs de Jordan

$$B_\lambda = \begin{pmatrix} \lambda & 1 & \ldots & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}$$

et il suffit de connaître e^{B_λ}. Or $B_\lambda = \lambda I + N$, où N est nilpotente d'ordre p et où p est la taille du bloc de Jordan. La matrice N commute avec chaque matrice diagonale. Donc $e^{B_\lambda} = e^\lambda e^N$ et $e^N = I + N + \frac{1}{2!}N^2 + \dots \frac{1}{(p-1)!}N^{p-1}$. Explicitement, on trouve (en introduisant le paramètre t) :

$$e^{tN} = U_p(t) = \begin{pmatrix} 1 & t & \dots & \frac{1}{(p-1)!}t^{p-1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Conclusion : avec P la matrice ayant comme colonnes les vecteurs qui donnent une forme de Jordan avec les blocs de Jordan B_{λ_i} de taille p_i, $i = 1, \dots, q$ on a

$$e^{tA} = P \begin{pmatrix} e^{\lambda_1 t}U_{p_1}(t) & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t}U_{p_2}(t) & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & e^{\lambda_m t}U_{p_q}(t) \end{pmatrix} P^{-1}.$$

Conclusion :

2.2. Théorème. *La solution unique du problème de Cauchy $y' = Ay$, $y(t_0) = y_0$ est donnée par $y(t) = e^{(t-t_0)A}y_0$, où $e^{(t-t_0)A}$ est la matrice qu'on peut calculer lors de la formule précédente.*

2.3. Remarque. Souvent il est utile de calculer e^A directement. Par exemple, avec

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

on a

$$e^{Jt} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

On peut utiliser cela pour trouver la solution d'un problème qui vient de la physique : déterminer la trajectoire d'un particule de masse m et de charge électrique q sous l'influence d'un champ magnétique $\vec{B}$. Les lois de Newton et de Lorentz donnent l'équation différentielle pour la vitesse $\vec{v}$:

$$\frac{d\vec{v}}{dt} = \left(\frac{q}{m}\right) \vec{v} \wedge \vec{B}.$$

Pour résoudre cette équation on prend une base orthonormée $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ telle que $\vec{B} = B\vec{e}_3$. L'équation se transforme en

$$\frac{d\vec{v}}{dt} = (qB/m) \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{v}.$$

avec solution ($\omega = (qB/m)$) :

$$\begin{pmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{pmatrix} = \begin{pmatrix} \cos(\omega t) & -\sin(\omega t) & 0 \\ \sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1(0) \\ v_2(0) \\ v_3(0) \end{pmatrix}.$$

Intégrant une fois de plus, on trouve :

$$\vec{x(t)} - \vec{x(0)} - \begin{pmatrix} -v_2(0)/\omega \\ v_1(0)/\omega \\ 0 \end{pmatrix} = \begin{pmatrix} \sin(\omega t) & \cos(\omega t) & 0 \\ -\cos(\omega t) & \sin(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1(0)/\omega \\ v_2(0)/\omega \\ 0 \end{pmatrix} + tv_3(0)\vec{e}_3$$

et donc $\vec{x}(t)$ décrit une hélice de rayon $\frac{1}{\omega}\sqrt{v_1(0)^2 + v_2(0)^2}$ et le mouvement est "uniforme".

§ 3. Systèmes linéaires non-homogènes à coefficients constants

Une méthode générale pour trouver une solution particulière est celle de la variation des constantes. On essaie $Y(t) = e^{tA}V(t)$ et on trouve

$$\begin{aligned} Y' =& Ae^{tA}V + e^{tA}V' \\ =& AY + e^{tA}V' = \\ =& AY + B \end{aligned}$$

donc $V' = e^{-tA}B$ donc $V(t) = \int_{t_0}^{t} e^{-sA}B(s)ds$ donne une solution avec $Y(t_0) = 0$. On trouve comme solution avec $Y(t_0) = Y_0$:

$$Y(t) = e^{(t-t_0)A}(Y_0) + \int_{t_0}^{t} e^{(t-s)A}B(s)ds.$$

Souvent c'est plus facile d'essayer de trouver la solution d'une autre façon.

3.1. Exemple. On reprend le dernier exemple du paragraphe précédent en rajoutant un champ électrique $\vec{E}$. La loi du mouvement devient :

$$\frac{d\vec{v}}{dt} = \left(\frac{q}{m}\right)\vec{v} \wedge \vec{B} + \frac{q}{m}\vec{E}.$$

Si $\vec{E}//\vec{B}$ alors on a la solution évidente $\vec{v}(t) = \dfrac{q}{m}t\vec{E}$ et donc

$$\vec{x}(t) - \vec{x}(0) - \begin{pmatrix} -v_2(0)/\omega \\ v_1(0)/\omega \\ 0 \end{pmatrix} = \begin{pmatrix} \sin(\omega t) & \cos(\omega t) & 0 \\ -\cos(\omega t) & \sin(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1(0)/\omega \\ v_2(0)/\omega \\ 0 \end{pmatrix} + tv_3(0)\vec{e}_3 + \frac{q}{2m}t^2\vec{E}.$$

Le mouvement est hélicoïdal mais accéléré.

Si $\vec{E}$ et $\vec{B}$ ne sont pas parallèles, on décompose

$$\vec{E} = \vec{E}_{//} + \vec{E}_{\perp}$$

et, avec $\vec{u}$ tel que

$$\vec{u} \wedge \vec{B} = \vec{E}_{\perp}$$

on trouve l'équation

$$\frac{d\vec{v}}{dt} = \left(\frac{q}{m}\right)\vec{v} + \vec{u} \wedge \vec{B} + \frac{q}{m}\vec{E}_{//}$$

et donc il faut remplacer $\vec{v}$ par $\vec{v} + \vec{u}$ ce qui donne la solution

$$\vec{x}(t) - \vec{x}(0) - \begin{pmatrix} -v_2(0)/\omega \\ v_1(0)/\omega \\ 0 \end{pmatrix} = \begin{pmatrix} \sin(\omega t) & \cos(\omega t) & 0 \\ -\cos(\omega t) & \sin(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1(0)/\omega \\ v_2(0)/\omega \\ 0 \end{pmatrix} + tv_3(0)\vec{e}_3 - t\vec{u} + \frac{q}{2m}t^2\vec{E}_{//}.$$

Le mouvement est hélicoïdal mais cette fois spiralisant autour d'un paraboloïde.

§ 4. Résolvante

On revient au cas des systèmes homogènes à coefficients variables $Y' = A(t)Y$, $A(t) \in M_m$ une fonction continue sur $J \subset \mathbb{R}$ et on a l'espace vectoriel des solutions

$$\mathrm{Sol} = \{Y \in \mathcal{C}^1(J, \mathbb{R}^m); Y' = A(t)Y\}$$

et pour n'importe quel $s \in J$ l'isomorphisme

$$\begin{aligned} \Phi_s \ : \mathrm{Sol} &\to \mathbb{R}^m \\ Y &\mapsto Y(s) \end{aligned}.$$

4.1. Définition. *$\forall t, t' \in J$ la résolvante $R(t,t')\ : \mathbb{R}^m \to \mathbb{R}^m$ est définie comme $R(t,t') = \Phi_t \circ \Phi_{t'}^{-1}$, et donc $R(t,t')$ associe à V la valeur en t de l'unique solution Y t.q. $Y(t') = V$.*

Remarque. Si on peut calculer directement la résolvante $R(t,t')$ on trouve aussi les solutions, car $R(t,t_0)V$ est l'unique solution $Y(t)$ telle que $Y(t_0) = V$. Le lemme suivant dit que la résolvante satisfait une équation différentielle semblable à l'équation originale. Donc il y a peu de chance qu'il soit plus facile de calculer la résolvante. Elle est surtout d'un intérêt théorique.

4.2. Lemme.

1. $\dfrac{dR(t,t')}{dt} = A(t)R(t,t')$,
2. $R(t,t) = I$,
3. $R(t'',t') \circ R(t',t) = R(t'',t)$.

Démonstration.

1. On a
$$\left(\frac{dR(t,t')}{dt}\right) V = \frac{d\,(R(t,t')V)}{dt} = \frac{dY(t)}{dt} = AY(t) = AR(t,t')V.$$

Les autres propriétés sont claires. ■

4.3. Exemple. Supposons que $A(s)$ et $A(t)$ commutent $\forall s, t \in J$. Alors

$$R(t,t') = \exp\left(\int_{t'}^{t} A(s)ds\right).$$

Pour montrer cela, il suffit de montrer que $M(t) = \exp\left(\int_{t'}^{t} A(s)ds\right)$ est la solution de $M' = AM$ avec $M(t') = I$. Or, de $A(s)A(s') = A(s')A(s)$ on tire que $\int_a^b A(u)du$ et $\int_c^d A(u)du$ commutent et donc $M(t+h) = \exp\left(\int_{t'}^{t} A(s)ds + \int_t^{t+h} A(s)ds\right) = \exp\left(\int_t^{t+h} A(s)ds\right) \cdot \exp\left(\int_{t'}^{t} A(s)ds\right) = \exp\left(\int_t^{t+h} A(s)ds\right) M(t)$ et puisque $\int_t^{t+h} A(s)ds = hA(t) + O(h^2)$ on trouve

$$M(t+h) = \left(I + h(A(t) + O(h^2)\right) M(t) = M(t) + hA(t)M(t) + O(h^2)$$

et donc

$$\lim_{h \to 0} \frac{M(t+h) - M(t)}{h} = A(t)M(t).$$

Finalement il est clair que $M(t') = I$.

Cas particulier : $A(t) = f(t)U + g(t)V$ avec $UV = VU$. Par exemple $A(t) = \begin{pmatrix} a(t) & -b(t) \\ b(t) & a(t) \end{pmatrix}$ donne

$$R(t,t') = \exp\left(\int_{t'}^{t} a(s)ds\right) \begin{pmatrix} \cos\left(\int_{t'}^{t} b(s)ds\right) & -\sin\left(\int_{t'}^{t} b(s)ds\right) \\ \sin\left(\int_{t'}^{t} b(s)ds\right) & \cos\left(\int_{t'}^{t} b(s)ds\right) \end{pmatrix}.$$

Utilisant la résolvante on peut aussi trouver une solution de l'équation inhomogène $Y' = A(t)Y + B(t)$. On cherche $Y = R(t,t')V(t)$ telle que $Y(t) = V$. Je montrerai :

$$Y = R(t,t')V + \int_{t'}^{t} R(t,s)B(s)ds.$$

Or, on a

$$\begin{aligned} \frac{dY}{dt} =& \frac{dR(t,t')}{dt}V(t) + R(t,t')\frac{dV}{dt} \\ =& A(t)R(t,t')V(t) + R(t,t')\frac{dV}{dt} \\ =& A(t)Y(t) + R(t,t')\frac{dV}{dt} \\ =& A(t)Y(t) + B(t) \end{aligned}$$

et donc

$$\frac{dV}{dt} = \int_{t'}^{t} R(t',s)B(s)ds$$

et

$$Y = R(t,t')V(t) = \int_{t'}^{t} R(t,t')R(t',s)B(s)ds = \int_{t'}^{t} R(t,s)B(s)ds$$

est la solution avec $Y(t') = 0$. On obtient la solution avec $Y(t) = V$ en rajoutant $R(t,t')V$.

§ 5. Stabilité, singularités des champs vectoriels

Regardons la situation générale d'une équation $y' = f(t,y)$ avec une solution $y(t,z)$ telle que $y(t_0) = z$ définie sur $[t_0, \infty[$. On dit que cette solution est *stable* resp. *asymptotiquement stable* si chaque solution $y(t,z')$ pour z' près de $y(t,z)$ reste près de z resp. converge vers $y(t,z)$.

Précisément, stabilité veut dire qu'il y a une boule B autour de z et une constante $C > 0$ telles que $\forall z' \in B$, $||y(t,z) - y(t,z')|| \leq C||z - z'||$, et $y(t,z)$ est asymptotiquement stable si de plus on peut trouver une fonction $c : [t_0, \infty[\to \mathbb{R}_+$ telle que $\lim_{t\to\infty} c(t) = 0$ et $||y(t,z) - y'(t,z')|| \leq c(t)||z - z'||$. Une solution est *instable* si elle n'est pas stable.

Dans le cas d'un système linéaire homogène à coefficients constants $Y' = AY$, les solutions sont $y(t,z) = e^{(t-t_0)A}z$ ce qui donne :

Les solutions sont stables	$\Leftrightarrow$	$\lim_{t\to\infty} \|\|e^{(t-t_0)A}\|\|_\infty \leq$ Constante
Et asymptotiquement stables	$\Leftrightarrow$	$\lim_{t\to\infty} \|\|e^{(t-t_0)A}\|\|_\infty = 0$

Si A est diagonalisable avec valeurs propres $\lambda_j = a_j + \mathrm{i}b_j$, $j = 1, \ldots, m$ on a

$$||e^{(t-t_0)A}||_\infty = \max_j e^{(t-t_0)a_j}$$

et donc les solutions sont stables si et seulement si $a_j \leq 0$, $j = 0, \dots, m$, asymptotiquement stable si et seulement si $a_j < 0$, $j = 1, \dots, m$.

Si A n'est pas diagonalisable on se ramène au cas d'un seul bloc de Jordan : $A = \lambda I + N$, $\lambda = a + b\mathrm{i}$. Alors

$$||e^{(t-t_0)A}||_\infty = e^{(t-t_0)a} \cdot ||\sum_k \frac{1}{k!}(t-t_0)^k N^k||_\infty.$$

Puisque $||\sum_k \frac{1}{k!}(t-t_0)^k N^k||_\infty$ est un polynôme en t, si $a < 0$ chaque solution est asymptotiquement stable, mais si $a > 0$ l'exponentielle l'emporte sur le polynôme et les solutions ne sont pas stables. Pour $a = 0$, le polynôme n'étant pas constant, les solutions sont aussi non stables. En conclusion :

5.1. Théorème. *Les solutions de $Y' = AY$ sont asymptotiquement stables si et seulement si la partie réelle de chaque valeur propre est négative. Elles sont stables si la partie réelle de chaque valeur propre avec un ou plusieurs blocs de Jordan de taille ≥ 2 est négative et si de plus la partie réelle d'une valeur propre avec un bloc diagonale est ≤ 0.*

On retourne d'abord au cas général $Y' = f(Y)$ avec $f(Y)$ indépendant de t et dérivable en Y. Cette fonction définit un champ vectoriel $Y \mapsto f(Y)$ dans $\mathbb{R}^m$. Si $f(Y_0) \neq 0$ on a une *point non-singulier* du champs. Sinon Y_0 est singulier. Autour d'un point non-singulier les trajectoires $Y(t)$ sont presque parallèles à $Y(0)$.

On regarde en détail le cas linéaire de dimension 2, donc le cas d'un champ linéaire AY de vecteurs avec A matrice à valeurs réelles. On suppose que A soit inversible et donc 0 est le seul point singulier. Classifions les points singuliers :

1. Deux valeurs propres réelles inégales λ_1, λ_2. Après changement de base les solutions sont

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 e^{\lambda_1 t} \\ y_0 e^{\lambda_2 t} \end{pmatrix}$$

   avec courbes intégrales

$$y = C|x|^{\frac{\lambda_2}{\lambda_1}}, \quad C \in \mathbb{R}.$$

   Trois cas possibles :

   A. $0 < \lambda_1 < \lambda_2$: on dit que c'est un nœud impropre instable.

   B. $\lambda_2 < \lambda_1 < 0$: un nœud impropre stable.

   C. $\lambda_1 < 0 < \lambda_2$: un col instable.

2. $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$. Deux cas :

   A. Le cas diagonalisable : les solutions sont des droites et on a (par définition) un nœud stable si $\lambda < 0$ et un nœud instable si $\lambda > 0$.

   B. Le cas $\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$ avec solution :

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 e^{\lambda t} \\ (y_0 + x_0 t) e^{\lambda t} \end{pmatrix}.$$

   Puisque chaque courbe de solution passe la droite $|x_0| = 1$, pour trouver une trajectoire différente de $x = 0$, on peut choisir $x_0 = \pm 1$ et on trouve

$$y = y_0|x| + \frac{|x|}{\lambda} \log|x|.$$

Pour $\lambda > 0$ on trouve un nœud exceptionnel instable et pour $\lambda < 0$ un nœud exceptionnel stable.

3. Les valeurs propres sont complexes, disons $a \pm b\mathrm{i}$. On peut trouver une base de $\mathbb{R}^2$ telle que

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

On écrit $z = x + \mathrm{i}y$ et on trouve

$$z(t) = z_0 e^{(a+b\mathrm{i})t}.$$

En coordonnées polaires $z = re^{\mathrm{i}\theta}$ on trouve

$$\begin{aligned} r &= r_0 e^{at} \\ \theta &= \theta_0 + bt \end{aligned}$$

et donc

$$r = r_0 e^{\frac{a}{b}(\theta - \theta_0)}$$

un spirale logarithmique si $a \neq 0$ et un cercle pour $a = 0$. Pour $a > 0$ on a un foyer instable et pour $a < 0$ un foyer stable.

Chapitre 5. Équations différentielles linéaires d'ordre quelconque

On applique la théorie du chapitre précédent au cas d'un système provenant d'une équation linéaire d'ordre quelconque. La théorie se simplifie. La notion centrale dans le cas des coefficients constants est celle du polynôme caractéristique associé obtenu en remplaçant d^k/dt^k par λ^k.

§ 1. Le cas général

La forme générale d'une équation linéaire d'ordre p est :

$$\text{(lin)} \qquad a_p(t)y^{(p)}(t) + \cdots + a_1(t)y'(t) + a_0(t)y_0(t) + b(t) = 0, \quad t \in J.$$

On a déjà vu que cette équation est équivalente à un système, qui dans ce cas s'écrit :

$$\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(p-2)} \\ y^{(p-1)} \end{pmatrix}' = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \cdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{a_0}{a_p} & -\frac{a_1}{a_p} & \cdots & \cdots & -\frac{a_{p-1}}{a_p} \end{pmatrix} \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(p-2)} \\ y^{(p-1)} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\frac{b}{a_p} \end{pmatrix}$$

Le théorème 1.1 du Chapitre 4 donne :

1.1. Théorème. *Il y a une unique solution maximale de* (lin), *valable sur J qui satisfait les conditions initiales $y(t_0) = y_0$, $y'(t_0) = y_1, \ldots y^{(p-1)}(t_0) = y_{p-1}$ ($t_0 \in J$).*

Soit Sol *l'ensemble des solutions de l'équation homogène associée. Alors* Sol *est un espace vectoriel de dimension p et l'application $y \mapsto (y(t_0), \ldots, y^{(p-1)}(t_0))$ donne un isomorphisme de* Sol *sur $\mathbb{R}^p$.*

Dans le cas inhomogène, les solutions forment un espace affine de la forme solution particulière + Sol.

§ 2. Coefficients constants sans second membre

Il s'agit de résoudre

$$\text{(LC)} \qquad a_p y^{(p)}(t) + \cdots + a_1 y'(t) + a_0 y_0(t) = 0, \quad t \in \mathbb{R}.$$

On calcule d'abord le polynôme caractéristique de A, matrice associée :

$$\det(A - \lambda I) = \frac{(-1)^{p+1}}{a_p}(a_p\lambda^p + a_{p-1}\lambda^{p-1} + \cdots a_0)$$

et donc les valeurs propres sont les racines $\lambda_1, \ldots, \lambda_s$ du *polynôme caractéristique* :

$$a_p\lambda^p(t) + \cdots + a_1\lambda + a_0 = 0.$$

2.1. Théorème. Sol, *l'espace des solutions de* (LC)*, admet une base qui est formée des fonctions :*

$$f_{q,j} = t^q e^{\lambda_j t}, \quad j = 1, \dots, s, \quad q = 0, \dots, p_j - 1, \quad p_j = \text{multiplicité de } \lambda_j.$$

Démonstration. La théorie générale dit qu'une solution du système $Y' = AY$ est un vecteur avec chaque coordonnée combinaison linéaire des fonctions $f_{q,j}$. En particulier $y(t)$, la première coordonnée est donc une telle combinaison. Il y a précisément $\sum p_j = p$ fonctions $f_{q,j}$. Si chaque fonction $f_{q,j}$ elle-même est solution de (LC) on a donc forcément une base. Pour vérifier cela, on introduit

$$P(\frac{d}{dt}) = \sum_{i=1}^{p} a_i \frac{d^i}{dt^i}$$

et on calcule :

$$\begin{aligned}
P(\frac{d}{dt})(t^q e^{\lambda t}) &= P(\frac{d}{dt}) \left(\frac{d^q}{d\lambda^q} e^{\lambda t} \right) \\
&= \frac{d^q}{d\lambda^q} \left(P(\frac{d}{dt}) e^{\lambda t} \right) \\
&= \frac{d^q}{d\lambda^q} \left(P(\lambda) e^{\lambda t} \right) \\
&= 0
\end{aligned}$$

si λ est racine de $P(\lambda)$. ∎

Si les racines sont complexes il faut travailler sur $\mathbb{C}$, mais si l'équation a des coefficients réels les racines complexes se groupent par paires, une paire consistant en un nombre et son complexe conjugué : $a + b\mathrm{i}$ et $a - b\mathrm{i}$. Dans ce cas, pour trouver une base qui est formée de fonctions réelles on peut remplacer $e^{(a+b\mathrm{i})t}$ et $e^{(a-b\mathrm{i})t}$ par $e^{at}\cos(bt)$ et $e^{at}\sin(bt)$.

2.2. Exemples.

1. $y'' + 4y = 0$. Ici $P(\lambda) = (\lambda + 2\mathrm{i})(\lambda - 2\mathrm{i})$. Une base de solutions sur $\mathbb{R}$ est donnée par $\{\cos(2t), \sin(2t)\}$.
2. $y'' - 4y' + 4y = 0$. Ici $P(\lambda) = (\lambda - 2)^2$ et donc Sol a une base $\{e^{2t}, te^{2t}\}$.

§ 3. Coefficients constants avec second membre

La méthode générale (variation des constantes) fonctionne ici, mais souvent il y a d'autre façons de résoudre un système homogène.

3.1. Exemples.

1. Si $b(t)$ est un polynôme de degré d on peut essayer un polynôme de degré d. Par exemple

$$2y'' + y' + 3y = t^3 + t.$$

On essaye $y = b_3 t^3 + b_2 t^2 + b_1 t + b_0$ et la substitution donne un système d'équations linéaires, qu'on peut résoudre. On trouve :

$$y = \frac{1}{3}t^3 - \frac{1}{3}t^2 - \frac{7}{9}t + \frac{19}{27}.$$

2. Si $b = e^{at}$ tel que $P(a) \neq 0$ on trouve la solution particulière $\dfrac{1}{P(a)} e^{at}$ car $P(d/dt)e^{at} = P(a)e^{at}$.

Maintenant la méthode des variations des constantes : supposons que $\{v_1, \ldots, v_p\}$ est une base de solutions de $a_p y^{(p)} + \cdots a_0 y = 0$. On cherche une combinaison linéaire de

$$V_i = \begin{pmatrix} v_i \\ v_i' \\ \vdots \\ v_i^{(p-1)} \end{pmatrix}, \quad i = 1, \ldots, p$$

telle $Y = b_1 V_1 + \ldots + b_p V_p$ soit solution de $Y' = AY + B$. La substitution donne $\sum_i b_i' V_i = B$, une équation linéaire en b_i' avec matrice inversible

$$U = \begin{pmatrix} v_1 & \cdots & v_p \\ \vdots & \cdots & \vdots \\ v_1^{(p-1)} & \cdots & v_p^{(p-1)}. \end{pmatrix}$$

Cette équation donne les b_i' et donc les b_i.

3.2. Exemple. $y'' + 4y = \tan(t), \quad t \in]-\pi/2, \pi/2[$. On trouve

$$U = \begin{pmatrix} \cos 2t & \sin 2t \\ -2\sin 2t & 2\cos 2t \end{pmatrix}$$

et

$$\begin{aligned} b_1' &= -\frac{1}{2} \sin 2t \tan t = -\sin^2 t \\ b_2' &= \frac{1}{2} \tan(t) \cos 2t \end{aligned}$$

avec solution

$$\begin{aligned} b_1 &= -\frac{1}{2} t + \frac{1}{4} \sin 2t \\ b_2 &= -\frac{1}{4} \cos 2t + \frac{1}{2} \log(\cos t) \end{aligned}$$

et donc la solution particulière $-\frac{1}{2} t \cos 2t + \frac{1}{2} \sin 2t \log(\cos t)$.