



EQUIPE D'ANALYSE

ESA 7064 – CNRS

Université Pierre et Marie Curie - Paris 6

Tour 46 - 0 - Boîte 186 - 4, place Jussieu - 75252 PARIS CEDEX 05
Tél : (33-1) 44 27 53 49 - Télécopie : (33-1) 44 27 25 55 - lana@ccr.jussieu.fr

EQUATIONS DIFFERENTIELLES METHODES DE RESOLUTION NUMERIQUE

Licence B

Approximation numériques des fonctions,
des intégrales, des solutions d'équations.

Equations différentielles :

Approximation numérique des solutions.

Sylvie DELABRIERE

2002-2003

1

RAPPEL DES RÉSULTATS FONDAMENTAUX DE CALCUL DIFFÉRENTIEL

1.1 Différentiabilité

Définition 1.1.1.

1) Soit E et F deux espaces vectoriels normés, $a \in E$ et V un voisinage de a dans E . Une application $f : V \rightarrow F$ est différentiable en a s'il existe une application linéaire continue de E dans F , notée $Df(a)$, telle

$$\forall h \in E \text{ tel que } a + h \in V, f(a + h) = f(a) + Df(a)h + \|h\| \varepsilon(h)$$

où $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ dans F .

2) L'application linéaire continue $Df(a)$ de E dans F est appelée la différentielle de f en a ou encore dans certains cas l'application linéaire tangente à f en a .

3) L'application affine $x \rightarrow f(a) + Df(a)(x - a)$ de E dans F , est appelée l'application affine tangente à f en a .

Remarque 1.1.2. On notera qu'une application différentiable en un point est en particulier continue en ce point.

Exemple 1.1.3. Cas des fonctions d'une variable réelle : une fonction $f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ d'une

variable réelle définie sur un voisinage V de $a \in \mathbb{R}$, à valeurs dans \mathbb{R}^m pour un $m \geq 1$, est différentiable en a si ses coordonnées sont dérivables en a et sa différentielle au point a est

l'application linéaire de \mathbb{R} dans \mathbb{R}^m représentée par le vecteur de \mathbb{R}^m , $Df(a)$, que l'on notera

$$\text{dans ce cas } f'(a) = \begin{pmatrix} f'_1(a) \\ f'_2(a) \\ \vdots \\ f'_m(a) \end{pmatrix}.$$

Théorème 1.1.4. (Dérivation suivant un vecteur).

Soit $a \in E$ et V un voisinage de a dans E . Si $f : V \rightarrow F$ est différentiable en a , alors quel que soit $h \in E$, on a

$$Df(a)h = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t}$$

DÉMONSTRATION : Soit $h \in E$ donné. On écrit la définition de la différentiabilité de f au point a , en utilisant le vecteur th :

$$f(a + th) = f(a) + tDf(a)h + |t| \|h\| \varepsilon(th) = f(a) + tDf(a)h + |t| \varepsilon_1(t)$$

où on a posé $\varepsilon_1(t) = \|h\| \varepsilon(th)$.

Comme $\varepsilon_1(t) \rightarrow 0$ quand $t \rightarrow 0$, ceci implique bien que

$$Df(a)h = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t}$$

■

Remarque 1.1.5. La dérivée en $a \in V$ d'une fonction $f : V \subset E \rightarrow F$ selon un vecteur h est une dérivée au sens usuel dans \mathbb{R} , précisément $Df(a)h = \varphi'(0)$ où $\varphi(t) = f(a + th)$ est une fonction de la variable réelle t .

Corollaire 1.1.6. Soit $a \in E$ et V un voisinage de a dans E . Si $f : V \rightarrow F$ est différentiable en a , alors, $Df(a)$ est unique.

DÉMONSTRATION : L'application linéaire $Df(a)$ est en effet entièrement déterminée par les égalités du théorème 1.1.4. ■

Définition 1.1.7. Lorsque $E = \mathbb{R}^m$, on notera $x = (x_1, x_2, \dots, x_m)$ les coordonnées d'un vecteur $x \in \mathbb{R}^m$. Alors si f est différentiable en $a \in V$, les dérivées en $a \in V$ de f selon les vecteurs de base de \mathbb{R}^m , $\{e_1, e_2, \dots, e_m\}$, sont appelées dérivées partielles de f en a et notées $\frac{\partial f}{\partial x_i}(a)$ pour $i = 1, 2, \dots, m$.

Par définition, on a donc pour $i = 1, 2, \dots, m$:

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t}$$

Définition 1.1.8. Lorsque $E = E_1 \times E_2$, on notera $x = (x_1, x_2)$ les coordonnées d'un vecteur $x \in E_1 \times E_2$. Alors si f est différentiable en $a \in V \subset E_1 \times E_2$, les différentielles partielles en $a = (a_1, a_2) \in V$ de f seront notées $D_{x_1}f(a)$ et $D_{x_2}f(a)$ et sont par définition les différentielles des fonctions partielles $f_1(x_1) = f(x_1, a_2)$ et $f_2(x_2) = f(a_1, x_2)$ en a_1 et a_2 respectivement.

On remarquera que si $E_1 = E_2 = \mathbb{R}$, les différentielles partielles de f coïncident avec les dérivées partielles définies ci-dessus.

Notons aussi que ces résultats se généralisent au produit d'un nombre quelconque d'espaces.

Proposition 1.1.9. 1) Toute application constante sur E est différentiable en tout point de E . Sa différentielle en tout point de E est la fonction nulle.

2) Soient E et F deux espaces vectoriels normés. Alors toute application linéaire continue $T : E \rightarrow F$ est différentiable en tout point $x \in E$ et sa différentielle en chaque point est l'application linéaire constante, égale à T .

3) Toute application affine continue définie par $\forall x \in E, Ax = b + Tx$, où $b \in F$ et $T \in \mathcal{L}(E; F)$ est différentiable en tout point $x \in E$ et sa différentielle en chaque point est l'application linéaire constante, égale à T .

4) Soient E et F deux espaces vectoriels normés. Alors toute application bilinéaire continue $B : E \times E \rightarrow F$ est différentiable en tout point $(x, y) \in E \times E$ et sa différentielle au point $(x, y) \in E \times E$ est l'application linéaire : $h \rightarrow B(x, h) + B(h, y)$.

DÉMONSTRATION : 1) est évident

2) Soit $x \in E$. On écrit :

$$T(x + h) - T(x) = T(h)$$

D'où $T'(x) = T$ pour tout $x \in E$.

3) se démontre comme 2).

4) On écrit :

$$B(x + h, y + k) - B(x, y) = B(x + h, y + k) - B(x + h, y) + B(x + h, y) - B(x, y) =$$

$$B(x + h, k) + B(h, y) = B(x, k) + B(h, y) + B(h, k)$$

Ceci donne le résultat puisque $\|B(h, k)\| \leq \|B\| \|h\| \|k\|$. ■

1.2 Théorème des Accroissements Finis

Définition 1.2.1. Soient E et F deux espaces vectoriels normés.

1) Une fonction f définie sur un ouvert Ω de E est dite différentiable sur Ω si elle est différentiable en tout point de Ω .

2) La fonction $Df : \Omega \rightarrow \mathcal{L}(E, F)$ est appelée la différentielle de f sur Ω .

3) Si f est différentiable sur Ω et si sa différentielle $f' : \Omega \rightarrow \mathcal{L}(E, F)$ est continue sur Ω on dira que f est continûment différentiable sur Ω . On dira aussi que f est de classe C^1 sur Ω .

Rappelons le résultat classique pour les fonctions d'une variable réelle, à valeurs réelles :

Théorème 1.2.2. (Théorème des Accroissements Finis scalaire)

Soit f une fonction continue d'un intervalle $[a, b]$, à valeurs dans \mathbb{R} , dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que :

$$f(b) - f(a) = f'(c)(b - a)$$

On en déduit une première extension pour les fonctions d'une variable vectorielle, à valeurs réelles :

Théorème 1.2.3. (Théorème des Accroissements Finis pour les fonctions d'une variable vectorielle à valeurs réelles)

Soit E un espace vectoriel normé et f une application différentiable d'un ouvert Ω de E , à valeurs dans \mathbb{R} . Si le segment $[a, a + h] \subset \Omega$, il existe $\theta \in]0, 1[$ tel que :

$$f(a + h) - f(a) = Df(a + \theta h)h$$

ce qui s'écrit aussi : il existe $c \in]a, a + h[$ tel que :

$$f(a + h) - f(a) = Df(c)h$$

Remarque 1.2.4. Le résultat précédent est faux si la fonction f est à valeurs dans un espace vectoriel de dimension supérieure ou égale à 2.

On peut à partir de ces résultats obtenir le théorème des Accroissements Finis pour les fonctions d'une variable vectorielle, à valeurs vectorielles :

Théorème 1.2.5. (Théorème des Accroissements Finis)

Soient E et F deux espaces vectoriels normés et f une application différentiable dans un ouvert Ω de E , à valeurs dans F . Si $[a, a + h] \subset \Omega$, on a :

$$\|f(a + h) - f(a)\| \leq \|h\| \sup_{x \in]a, a+h[} \|Df(x)\|$$

1.3 Théorèmes des fonctions implicites

Théorème 1.3.1. (Théorème des fonctions implicites)

Soient E un espace topologique, F et G deux espaces de banach et f une application d'un ouvert Ω de $E \times F$, à valeurs dans G .

Soient $a, b \in E \times F$ tels que $f(a, b) = 0$. On suppose de plus que f est continue sur Ω , que la différentielle partielle Df_y existe et est continue sur Ω et que $Df_y(a, b)$ est un isomorphisme de F sur G .

Alors, il existe un voisinage A de a dans E , un voisinage B de b dans F et une fonction continue g de A dans B , tels que la restriction à $A \times B$ de la relation $f(x, y) = 0$ soit équivalente à la relation : $y = g(x)$.

Théorème 1.3.2. (Différentiabilité des fonctions implicites)

Soient E, F et G trois espaces de banach et f une application continue d'un ouvert Ω de $E \times F$, à valeurs dans G .

On suppose qu'il existe un ouvert A de E , un ouvert B de F et une fonction continue g de A dans B , tels que la restriction à $A \times B$ de la relation $f(x, y) = 0$ soit équivalente à la relation : $y = g(x)$.

On suppose de plus que f est différentiable en $(a, b) \in A \times B$ et que $Df_y(a, b)$ est un isomorphisme de F sur G .

Alors, g est différentiable en a et

$$g'(a) = -(Df_y(a, b))^{-1} \circ Df_x(a, b)$$

Théorème 1.3.3. (Continuité des différentielles de fonctions implicites 1)

Soient E , F et G trois espaces de Banach et f une fonction continûment différentiable d'un ouvert Ω de $E \times F$ à valeurs dans G .

On suppose que la relation $f(x, y) = 0$ est équivalente sur le produit $A \times B$ de deux ouverts de E et F à la relation $y = g(x)$, où g est continue sur A et que pour tout $x, y \in A \times B$, $Df_y(x, y)$ est un isomorphisme de F sur G .

Alors g est continûment différentiable sur A .

Théorème 1.3.4. (Continuité des différentielles de fonctions implicites 2)

Soient E , F et G trois espaces de Banach et f une fonction continûment différentiable d'un ouvert Ω de $E \times F$ à valeurs dans G .

On suppose que $f(a, b) = 0$ et $Df_y(a, b)$ est un isomorphisme de F sur G .

Alors il existe un voisinage ouvert A de a dans E , un voisinage ouvert B de b dans F tels que la restriction à $A \times B$ de la relation $f(x, y) = 0$ soit équivalente à une relation $y = g(x)$, où g est continûment différentiable sur A .

1.4 Théorème d'inversion locale

Définition 1.4.1. Soient E et F deux espaces de Banach. On appelle difféomorphisme, respectivement difféomorphisme de classe C^1 , d'une partie ouverte A de E dans une partie ouverte B de F toute bijection f , de A sur B , différentiable de A dans B , respectivement continûment différentiable de A dans B , ainsi que son inverse f^{-1} .

Théorème 1.4.2. Soient E et F deux espaces de Banach et f une application de classe C^1 sur un ouvert de E , à valeurs dans F . Si $Df(a)$ est un isomorphisme de E sur F , il existe un voisinage ouvert A de a dans E et un voisinage ouvert B de $b = f(a)$ dans F tels que la restriction de f à A soit un difféomorphisme de classe C^1 de A sur B .

DÉMONSTRATION : On applique le théorème 1.3.3 à la relation $y - f(x) = 0$: il existe un voisinage ouvert A de a et un voisinage ouvert B de $b = f(a)$ tels que sur $A \times B$, la relation $y = f(x)$ soit équivalent à une relation $x = g(y)$ où g est continûment différentiable de B sur A .

Cela signifie que la restriction de f à A est un difféomorphisme de classe C^1 . ■

1.5 Notations différentielles

Dans la pratique, on utilisera des notations très suggestives, qui ont le très grand mérite de résumer les formules habituelles du calcul différentiel.

Si f est une fonction différentiable de m variables x_1, x_2, \dots, x_m , la différentielle de f sera notée df et on définit alors la relation suivante :

$$df(x_1, x_2, \dots, x_n) = \frac{\partial f}{\partial x_1}(x_1, x_2, \dots, x_n)dx_1 + \frac{\partial f}{\partial x_2}(x_1, x_2, \dots, x_n)dx_2 + \dots + \frac{\partial f}{\partial x_m}(x_1, x_2, \dots, x_n)dx_m$$

En particuliers, si $x, y \in \mathbb{R}$ et si $y - f(x) = 0$, on aura par convention :

$$dy = df(x)dx$$

soit encore si $x \in \mathbb{R}$:

$$df(x) = \frac{dy}{dx}$$

Cette convention de notation permet de retrouver des règles de différentiation :

Application 1.5.1. Règle de différentiation des fonctions composées.

Si $z = g(y)$ et $y = f(x)$, où f et g sont des fonctions différentiables, la dérivée de $g \circ f$ s'obtient en différentiant successivement les deux relations :

$$dz = dg(y)dy, \quad dy = df(x)dx$$

et en reportant la valeur de dy fournie par la deuxième formule dans la première formule :

$$dz = dg(y)df(x)dx = dg(f(x))df(x)dx$$

et on retrouve ainsi la formule habituelle.

Application 1.5.2. Règle de différentiation des fonctions implicites.

Soient E, F et G trois espaces de Banach et f une application de classe C^1 d'un ouvert Ω de $E \times F$, à valeurs dans G .

En différentiant la relation $f(x, y) = 0$, on trouve :

$$\frac{\partial}{\partial x}f(x, y)dx + \frac{\partial}{\partial y}f(x, y)dy = 0$$

Si pour $x = a$ et $y = b$, l'application linéaire $\frac{\partial}{\partial y}f(a, b)$ est un isomorphisme de F sur G , alors la relation $f(x, y) = 0$ définit implicitement une relation $y = g(x)$ dans un voisinage de (a, b) et la différentielle de g se calcule à partir de la différentielle de la formule précédente :

$$dy = -\left(\frac{\partial}{\partial y}f(x, y)\right)^{-1} \frac{\partial}{\partial x}f(x, y)dx$$

On retrouve ainsi la formule donnée au théorème 1.3.1.

2

MÉTHODES D'APPROXIMATION NUMÉRIQUE DES FONCTIONS

Dans toute la suite, on notera \mathcal{P}_n l'espace vectoriel des fonctions polynômes sur \mathbb{R} , à coefficients dans \mathbb{R} , de degré inférieur ou égal à n .

On rappelle que $\dim \mathcal{P}_n = n + 1$.

L'espace vectoriel des fonctions continues sur un intervalle $[a, b]$ à valeurs dans \mathbb{R} sera noté $\mathcal{C}([a, b])$ et cet espace sera muni de la norme uniforme définie par

$$\|f\| = \sup_{x \in [a, b]} |f(x)|$$

2.1 Approximation d'une dérivée

Définition 2.1.1. Soit f une fonction réelle ou complexe, définie dans un intervalle I de \mathbb{R} . Soit x_0 un point de I . On dit que f est dérivable en x_0 , de dérivée ℓ , si

$$\lim_{h \rightarrow 0, h \neq 0} \frac{f(x_0 + h) - f(x_0)}{h} = \ell.$$

Cette dérivée se note $f'(x_0)$ ou aussi $\frac{df}{dx}(x_0)$. C'est donc la limite du quotient différentiel. La variable h est l'accroissement, qui peut-être positif ou négatif. Lorsque la fonction f a une dérivée en chaque point de l'intervalle I , on dit que f est dérivable dans I , de dérivée f' . Dans ce cas, on peut itérer cette définition, c'est-à-dire, définir la dérivée de f' , qu'on appelle dérivée seconde de f et qu'on note f'' et définir ainsi de suite f''' , etc.

On connaît des formules qui permettent de dériver les fonctions usuelles. Selon la fonction, le calcul de la dérivée est plus ou moins compliqué, mais comme c'est un procédé systématique, il existe des langages de calcul formel qui permettent, dans la plupart des cas, de calculer

l'expression de la dérivée f' à partir de l'expression de la fonction f . On peut donc légitimement se demander à quoi sert d'approcher une dérivée.

Il y a deux cas importants où on ne sait pas calculer la dérivée de manière exacte et où on est obligé de l'approcher.

- Le premier cas se produit quand on connaît les valeurs que prend la fonction f en un certain nombre de points, mais que l'expression de f n'est pas connue. Par exemple, ces valeurs de f peuvent venir de mesures physiques ou d'un calcul précédent.

- Le second cas, qui est le plus fréquent, se produit quand la fonction f fait précisément partie des inconnues du problème. Par exemple, lorsque f est la solution d'une équation différentielle ou d'une équation aux dérivées partielles. C'est pourquoi il est important de savoir bien approcher une dérivée.

Dans la définition 2.1.1, le signe de h n'est pas précisé. Mais ici, nous adopterons la convention que h est strictement positif.

Définition 2.1.2. 1) Pour $h > 0$ et pour un point x_0 quelconque, on définit la différence finie en avant ou progressive de f au point x_0 :

$$\Delta f(x_0) = f(x_0 + h) - f(x_0)$$

2) la différence finie en arrière ou régressive de f au point x_0 :

$$\nabla f(x_0) = f(x_0) - f(x_0 - h)$$

3) et la différence finie centrée de f au point x_0 :

$$\delta f(x_0) = f\left(x_0 + \frac{h}{2}\right) - f\left(x_0 - \frac{h}{2}\right).$$

On remarque que ces définitions s'adaptent facilement à la somme de deux fonctions ; par exemple

$$\Delta(f + g)(x_0) = \Delta f(x_0) + \Delta g(x_0)$$

et au produit d'une fonction par une constante ; par exemple

$$\Delta(cf)(x_0) = c \Delta f(x_0)$$

On peut exprimer ces deux propriétés en disant que les trois opérateurs

$$f \rightarrow \Delta f(x_0) , f \rightarrow \nabla f(x_0) , f \rightarrow \delta f(x_0)$$

sont linéaires.

Grâce à ces propriétés, on peut facilement itérer ces formules, c'est-à-dire, définir les différences finies d'ordre deux, trois, etc. Par exemple, on pose par définition

$$\Delta^2 f(x_0) = \Delta(\Delta f(x_0))$$

Alors

$$\begin{aligned} \Delta(\Delta f(x_0)) &= \Delta(f(x_0 + h) - f(x_0)) = \Delta f(x_0 + h) - \Delta f(x_0) \\ &= f(x_0 + 2h) - f(x_0 + h) - \{f(x_0 + h) - f(x_0)\} \\ &= f(x_0 + 2h) - 2f(x_0 + h) + f(x_0). \end{aligned}$$

Donc

$$\Delta^2 f(x_0) = f(x_0 + 2h) - 2f(x_0 + h) + f(x_0)$$

Remarquons que les coefficients $(1, -2, 1)$ dans cette formule sont les mêmes que ceux du développement du binôme $(x - 1)^2$. Ce n'est pas une coïncidence. Pour s'en convaincre, il est commode de définir la translation de f au point x_0 :

$$E f(x_0) = f(x_0 + h)$$

On peut itérer cette définition en posant

$$E^2 f(x_0) = E(E f(x_0)) = E f(x_0 + h) = f(x_0 + 2h),$$

et de proche en proche, on obtient facilement la formule générale

$$E^n f(x_0) = f(x_0 + nh)$$

On peut aussi faire la translation inverse

$$E^{-1} f(x_0) = f(x_0 - h)$$

la demi translation

$$E^{\frac{1}{2}} f(x_0) = f(x_0 + \frac{h}{2})$$

et pas de translation du tout, opération notée I :

$$I f(x_0) = f(x_0)$$

Alors les formules de 2.1.2 s'écrivent à l'aide de translations :

$$\Delta f(x_0) = E f(x_0) - f(x_0) = (E - I)f(x_0)$$

$$\nabla f(x_0) = f(x_0) - E^{-1} f(x_0) = (I - E^{-1})f(x_0)$$

$$\delta f(x_0) = E^{\frac{1}{2}} f(x_0) - E^{-\frac{1}{2}} f(x_0) = (E^{\frac{1}{2}} - E^{-\frac{1}{2}})f(x_0)$$

Ces formules permettent de calculer facilement les puissances successives des différences finies et de comprendre pourquoi interviennent les coefficients du binôme. Par exemple

$$\begin{aligned} \Delta^4 f(x_0) &= (E - I)^4 f(x_0) = (E^4 - 4E^3 + 6E^2 - 4E + I)f(x_0) \\ &= f(x_0 + 4h) - 4f(x_0 + 3h) + 6f(x_0 + 2h) - 4f(x_0 + h) + f(x_0). \end{aligned}$$

Ici, on a utilisé le fait que $E I = I E = E$ et $I I = I$.

Définition 2.1.3. *Approximation de la dérivée première d'une fonction : à partir des différences finies 2.1.2, on peut définir trois approximations de $f'(x_0)$:*

$$\frac{\Delta f(x_0)}{h}, \quad \frac{\nabla f(x_0)}{h}, \quad \frac{\delta f(x_0)}{h}.$$

En comparant avec 2.1.1, on voit que h est destiné à tendre vers zéro. La variable h joue ici un rôle important :

Définition 2.1.4. h s'appelle le paramètre de discrétisation.

Il est facile de vérifier que si f est dérivable en x_0 , les deux premières approximations tendent bien vers $f'(x_0)$ quand h tend vers zéro. On verra que c'est le cas aussi de la troisième.

Mais, en calcul numérique, on ne se contente pas d'établir qu'une approximation converge : on veut aussi estimer sa vitesse de convergence. En effet, il y a des approximations dont on peut montrer qu'elles convergent, mais avec une telle lenteur qu'elles sont inutilisables en pratique.

Proposition 2.1.5. On appelle erreur dans l'approximation de la dérivée d'une fonction par $\frac{\Delta f(x_0)}{h}$ la quantité :

$$\left| \frac{\Delta f(x_0)}{h} - f'(x_0) \right|.$$

On a des définitions analogues avec les approximations $\frac{\nabla f(x_0)}{h}$ et $\frac{\delta f(x_0)}{h}$.

Prenons, pour fixer les idées, la première approximation de $f'(x_0)$. Pour estimer la vitesse avec laquelle elle converge vers $f'(x_0)$, on dispose d'un excellent outil mathématique, dont on fera ici un usage constant : la Formule de Taylor. Rappelons auparavant deux résultats fondamentaux du cours de mathématiques. Le premier est le Théorème de Rolle :

Théorème de Rolle 2.1.6. On suppose que f est dérivable en tout point de l'intervalle $[a, b]$ et que

$$f(a) = f(b).$$

Alors il existe (au moins) un point c tel que $a < c < b$ où

$$f'(c) = 0.$$

C'est-à-dire, il existe au moins un point où le graphe de f a une tangente horizontale.

Le second est le Théorème des Accroissements Finis, qui se déduit facilement du Théorème de Rolle.

Théorème des Accroissements Finis 2.1.7. On suppose que f est dérivable en tout point de l'intervalle $[a, b]$. Alors il existe (au moins) un point c tel que $a < c < b$ où on a l'égalité

$$f(b) - f(a) = (b - a) f'(c)$$

C'est-à-dire, il existe au moins un point où la tangente au graphe de f est parallèle à la corde.

Remarque 2.1.8. On ne doit appliquer ces deux théorèmes que si on est certain que la fonction est dérivable dans l'intervalle. Il est facile de construire des exemples de fonctions qui ne sont pas dérivables en un point de l'intervalle et pour lesquelles la conclusion de ces deux théorèmes est fausse.

La Formule de Taylor est une généralisation directe du Théorème des Accroissements Finis.

Théorème 2.1.9. (Formule de Taylor) On suppose que f est n fois dérivable en tout point de l'intervalle $[a, b]$. Alors il existe (au moins) un point c tel que $a < c < b$ où on a l'égalité

$$f(b) = f(a) + (b - a) f'(a) + \frac{(b - a)^2}{2} f''(a) + \frac{(b - a)^3}{6} f'''(a) + \dots \\ + \frac{(b - a)^{n-1}}{(n - 1)!} f^{(n-1)}(a) + \frac{(b - a)^n}{n!} f^{(n)}(c).$$

On voit bien que le Théorème des Accroissements Finis coïncide avec la Formule de Taylor à l'ordre $n = 1$.

Remarque 2.1.10. Comme pour les deux premiers théorèmes, il faut faire très attention à la dérivabilité des fonctions auxquelles on l'applique. On ne doit appliquer la Formule de Taylor d'ordre n que si on est certain que la fonction est n fois dérivable dans l'intervalle.

Proposition 2.1.11. Si f est 2 fois dérivable et si $|f''(c)| \leq M_2$ pour tout $c \in [x_0, x_0 + h]$, alors

$$\left| \frac{\Delta f(x_0)}{h} - f'(x_0) \right| \leq \frac{h}{2} M_2.$$

DÉMONSTRATION : Utilisons la Formule de Taylor d'ordre $n = 2$ pour majorer l'erreur dans 2.1.5 :

$$\left| \frac{\Delta f(x_0)}{h} - f'(x_0) \right| = \left| \frac{1}{h} [f(x_0 + h) - f(x_0)] - f'(x_0) \right| = \frac{h}{2} |f''(\xi)|,$$

où le point ξ est situé entre x_0 et $x_0 + h$: $x_0 < \xi < x_0 + h$.

Alors, on obtient bien :

$$\left| \frac{\Delta f(x_0)}{h} - f'(x_0) \right| \leq \frac{h}{2} M_2. \quad \blacksquare$$

Ceci veut dire que si la constante M_2 est de grandeur raisonnable, alors $\frac{\Delta f(x_0)}{h}$ tend vers $f'(x_0)$ avec la même vitesse que h .

Bien sûr, si la constante M_2 est très grande, il faut que h soit d'autant plus petit pour que l'erreur soit, elle aussi, petite. Evidemment, en théorie, on peut choisir h aussi petit qu'on veut, mais, en pratique il y a un seuil, dépendant du problème à traiter, en-dessous duquel, il n'est pas réaliste de choisir h . Donc, si f' a des trop grandes variations $\frac{\Delta f(x_0)}{h}$ n'est pas une bonne approximation de $f'(x_0)$, .

On aborde là une caractéristique essentielle du calcul numérique, qu'on retrouvera tout le temps : la précision d'une approximation dépend non seulement du type de l'approximation elle-même, mais aussi de la fonction à laquelle on l'applique. En calcul numérique, il n'y a aucune méthode d'approximation qui puisse être appliquée avec succès dans tous les cas.

Étudions maintenant l'approximation de la dérivée qui utilise une différence finie centrée ; c'est-à-dire, majorons l'erreur

$$\left| \frac{\delta f(x_0)}{h} - f'(x_0) \right|.$$

Proposition 2.1.12. Si f est 3 fois dérivable et si $|f'''(c)| \leq M_3$ pour tout $c \in [x_0 - \frac{h}{2}, x_0 + \frac{h}{2}]$, alors

$$\left| \frac{\delta f(x_0)}{h} - f'(x_0) \right| \leq \frac{1}{24} h^2 M_3.$$

DÉMONSTRATION : Avec la Formule de Taylor d'ordre $n = 2$, on a

$$\begin{aligned} f(x_0 + \frac{h}{2}) - f(x_0 - \frac{h}{2}) &= f(x_0) + \frac{h}{2} f'(x_0) + \frac{1}{2} \frac{h^2}{4} f''(\xi_1) \\ &\quad - \{ f(x_0) - \frac{h}{2} f'(x_0) + \frac{1}{2} \frac{h^2}{4} f''(\xi_2) \} \\ &= h f'(x_0) + \frac{h^2}{8} [f''(\xi_1) - f''(\xi_2)], \end{aligned}$$

où ξ_1 est un point compris entre x_0 et $x_0 + \frac{h}{2}$ et ξ_2 est un point compris entre $x_0 - \frac{h}{2}$ et x_0 . Remarquons que $f''(\xi_1)$ a le même coefficient que $f''(\xi_2)$, mais avec un signe opposé. Ceci suggère que si nous avons appliqué la Formule de Taylor d'ordre $n = 3$, ces deux termes se seraient éliminés. En effet, avec la Formule de Taylor d'ordre $n = 3$, on trouve

$$\begin{aligned} f(x_0 + \frac{h}{2}) - f(x_0 - \frac{h}{2}) &= f(x_0) + \frac{h}{2} f'(x_0) + \frac{1}{2} \frac{h^2}{4} f''(x_0) + \frac{1}{6} \frac{h^3}{8} f'''(\xi_3) \\ &\quad - [f(x_0) - \frac{h}{2} f'(x_0) + \frac{1}{2} \frac{h^2}{4} f''(x_0) - \frac{1}{6} \frac{h^3}{8} f'''(\xi_4)] \\ &= h f'(x_0) + \frac{h^3}{48} [f'''(\xi_3) + f'''(\xi_4)], \end{aligned}$$

où ξ_3 est un point compris entre x_0 et $x_0 + \frac{h}{2}$ et ξ_4 est un point compris entre $x_0 - \frac{h}{2}$ et x_0 , et ces points n'ont pas forcément de relation avec ξ_1 et ξ_2 .

Cette fois-ci, les coefficients de $f'''(\xi_3)$ et $f'''(\xi_4)$ ont le même signe ; donc ces deux termes ne s'élimineront pas si nous prenons une Formule de Taylor d'ordre plus élevé.

Pour conclure, on obtient :

$$\left| \frac{\delta f(x_0)}{h} - f'(x_0) \right| = \frac{1}{48} h^2 |f'''(\xi_3) + f'''(\xi_4)|.$$

Donc si la fonction f''' est bornée par

$$\forall x \in I, |f'''(x)| \leq M_3,$$

alors

$$\left| \frac{\delta f(x_0)}{h} - f'(x_0) \right| \leq \frac{1}{24} h^2 M_3.$$

■

Ceci veut dire que si la fonction f est trois fois dérivable et si sa dérivée troisième n'est pas trop grande, alors $\frac{\delta f(x_0)}{h}$ tend vers $f'(x_0)$ avec la même vitesse que h^2 . Comme h^2 tend vers zéro beaucoup plus vite que h , on conclut que, lorsque la constante M_3 n'est pas trop grande, $\frac{\delta f(x_0)}{h}$ est une bien meilleure approximation de $f'(x_0)$ que $\frac{\Delta f(x_0)}{h}$.

Remarque 2.1.13. Nous aurions gagné du temps en appliquant directement la Formule de Taylor d'ordre $n = 3$, mais comment savoir, avant de faire le calcul, que c'était cet ordre-là qui était le bon ? Malheureusement, c'est une question d'expérience ; on ne peut pas toujours le savoir, car il n'existe pas de règle qui permette de choisir a priori l'ordre de la Formule de Taylor qu'on doit appliquer.

Nous venons de voir que la différence finie centrée δ fournit une bonne approximation de la dérivée première. Ceci suggère d'utiliser δ^2 pour approcher la dérivée seconde :

Définition 2.1.14. L'approximation de la dérivée seconde d'une fonction f au point x_0 est :

$$\frac{\delta^2 f(x_0)}{h^2}.$$

Proposition 2.1.15. Soit f une fonction 4 fois dérivable telle que

$$\forall x \in [x_0 - \frac{h}{2}, x_0 + \frac{h}{2}], |f^{(4)}(x)| \leq M_4$$

L'erreur dans l'approximation de la dérivée seconde est majorée par :

$$|\frac{\delta^2 f(x_0)}{h^2} - f''(x_0)| \leq \frac{1}{12} h^2 M_4.$$

DÉMONSTRATION : La formule $\delta f(x_0) = E^{\frac{1}{2}} f(x_0) - E^{-\frac{1}{2}} f(x_0) = (E^{\frac{1}{2}} - E^{-\frac{1}{2}}) f(x_0)$ donne en itérant :

$$\begin{aligned} \delta^2 f(x_0) &= (E^{\frac{1}{2}} - E^{-\frac{1}{2}})^2 f(x_0) = (E - 2I + E^{-1}) f(x_0) \\ &= f(x_0 + h) - 2f(x_0) + f(x_0 - h). \end{aligned}$$

Pour majorer l'erreur

$$|\frac{\delta^2 f(x_0)}{h^2} - f''(x_0)|$$

nous allons utiliser la Formule de Taylor. Pour choisir l'ordre de cette formule, comparons avec le développement d'ordre $n = 3$ que nous avons utilisé ci-dessus. En remarquant que le dénominateur est ici h^2 (au lieu de h), si on espère obtenir une majoration en $O(h^2)$, il faut prendre la Formule de Taylor d'ordre au moins $n = 4$. Avec cette formule, on trouve

$$\begin{aligned} \delta^2 f(x_0) &= f(x_0) + hf'(x_0) + \frac{1}{2}h^2 f''(x_0) + \frac{1}{6}h^3 f'''(x_0) + \frac{1}{24}h^4 f^{(4)}(\xi_5) \\ &\quad - 2f(x_0) \\ &\quad + f(x_0) - hf'(x_0) + \frac{1}{2}h^2 f''(x_0) - \frac{1}{6}h^3 f'''(x_0) + \frac{1}{24}h^4 f^{(4)}(\xi_6) \\ &= h^2 f''(x_0) + \frac{1}{24}h^4 [f^{(4)}(\xi_5) + f^{(4)}(\xi_6)], \end{aligned}$$

où ξ_5 est un point compris entre x_0 et $x_0 + h$ et ξ_6 est un point compris entre $x_0 - h$ et x_0 . Remarquons que nous avons bien choisi d'emblée l'ordre de la Formule de Taylor, car les coefficients de $f^{(4)}(\xi_5)$ et $f^{(4)}(\xi_6)$ ont le même signe. Donc ces deux termes ne se seraient pas éliminés si nous avions pris une Formule de Taylor d'ordre plus élevé.

Pour conclure, nous trouvons

$$\left| \frac{\delta^2 f(x_0)}{h^2} - f''(x_0) \right| = \frac{1}{24} h^2 |f^{(4)}(\xi_5) + f^{(4)}(\xi_6)|.$$

Donc si la fonction $f^{(4)}$ est bornée, par exemple

$$\forall x \in I, |f^{(4)}(x)| \leq M_4,$$

alors

$$\left| \frac{\delta^2 f(x_0)}{h^2} - f''(x_0) \right| \leq \frac{1}{12} h^2 M_4.$$

Ceci veut dire que si la fonction f est quatre fois dérivable et si sa dérivée quatrième n'est pas trop grande, alors $\frac{\delta^2 f(x_0)}{h^2}$ tend vers $f''(x_0)$ avec la même vitesse que h^2 . ■

2.2 Interpolation d'une fonction par un polynôme

Définition 2.2.1. *Un polynôme est une fonction de la forme*

$$x \mapsto p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n, \quad c_n \neq 0.$$

Les quantités c_0, c_1, \dots, c_n sont les coefficients du polynôme. Ici, nous supposons que ce sont des nombres réels. L'entier n est le degré du polynôme. Enfin, la variable x est réelle, mais il est parfois nécessaire de considérer des cas où c'est une variable complexe, par exemple, quand on étudie les racines du polynôme.

Le but de ce paragraphe et suivants est d'étudier l'approximation d'une fonction par un polynôme. Il y a plusieurs façons de faire cette approximation ; ici nous étudierons un type particulier d'approximation polynomiale, appelé interpolation.

Définition 2.2.2. *L'interpolation d'une fonction f par un polynôme, consiste à remplacer la fonction f par un polynôme de degré n qui prend les mêmes valeurs que f en $n+1$ points. On appelle ces points, points d'interpolation et le polynôme, polynôme d'interpolation.*

Nous verrons dans les chapitres suivants que le polynôme d'interpolation rend des grands services. Par exemple, on sait facilement l'intégrer, alors qu'on ne sait pas en général intégrer la fonction qu'il interpole. On se sert également du polynôme d'interpolation quand on ne connaît que les valeurs de f en certains points.

2.3 Interpolation d'une fonction par un polynôme de degré 1

Le graphe d'un polynôme de degré 1 est une droite et une droite est déterminée par deux points distincts. Donc pour interpoler une fonction f , par un polynôme de degré 1, il faut connaître les valeurs de la fonction en 2 points distincts x_0 et x_1 . D'un point de vue géométrique, le graphe de ce polynôme d'interpolation est la droite qui passe par les points du plan de coordonnées $(x_0, f(x_0))$ et $(x_1, f(x_1))$. Pour trouver l'équation de cette droite, on applique le Théorème de Thalès, qui donne :

$$\frac{y - f(x_0)}{x - x_0} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Donc l'équation de la droite passant par ces deux points est

$$y = p_1(x) = f(x_0) + (x - x_0) \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

C'est bien un polynôme de degré 1 et il n'est défini que si $x_0 \neq x_1$.

Le facteur

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

est un quotient de la forme $\frac{\Delta f}{\Delta x}$. Plus précisément, si on écrit $x_1 = x_0 + h$, ce facteur se met sous la forme

$$\frac{f(x_0 + h) - f(x_0)}{h} \quad \text{c'est-à-dire} \quad \frac{\Delta f(x_0)}{h}.$$

Nous avons vu au chapitre précédent que ce quotient tend vers $f'(x_0)$ quand h tend vers zéro.

Définition 2.3.1. *Le quotient*

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

s'appelle différence divisée d'ordre 1 de f aux points x_0 et x_1 et se note $f[x_0, x_1]$.

Remarquons que, dans cette définition, l'ordre des points x_0 et x_1 est sans importance, car $f[x_0, x_1] = f[x_1, x_0]$. Si on définit la différence divisée d'ordre 0 de f au point x_0 par

$$f[x_0] = f(x_0)$$

on peut exprimer l'équation d'une droite passant par deux points entièrement avec des différences divisées :

$$y = p_1(x) = f[x_0] + f[x_0, x_1](x - x_0)$$

Définition 2.3.2. *Le polynôme $y = p_1(x) = f[x_0] + f[x_0, x_1](x - x_0)$ s'appelle le polynôme d'interpolation de Newton d'ordre 1 associé à f .*

Evidemment, on peut grouper les termes de ce polynôme d'interpolation de différentes façons ; par exemple, on peut aussi écrire

$$y = p_1(x) = f(x_0) \frac{x - x_1}{x_0 - x_1} + f(x_1) \frac{x - x_0}{x_1 - x_0}$$

Définissons les polynômes de degré un

$$l_0(x) = \frac{x - x_1}{x_0 - x_1} \quad , \quad l_1(x) = \frac{x - x_0}{x_1 - x_0} .$$

Alors on obtient :

$$y = p_1(x) = f(x_0)l_0(x) + f(x_1)l_1(x)$$

Remarquons que l_0 et l_1 ne dépendent que de x , x_0 et x_1 et pas de la fonction f ; en particulier,

$$l_0(x_0) = 1 \quad , \quad l_0(x_1) = 0 \quad , \quad l_1(x_0) = 0 \quad , \quad l_1(x_1) = 1 .$$

Définition 2.3.3. On dit que les polynômes l_0, l_1 sont les polynômes de base de degré 1 associés à x_0, x_1 .

Grâce à ces propriétés, si on considère le polynôme de degré 1,

$$\tilde{p}(x) = \alpha_0 l_0(x) + \alpha_1 l_1(x) ,$$

où α_0 et α_1 sont des constantes réelles, on peut dire, sans faire de calculs, que $\tilde{p}(x_0) = \alpha_0$ et $\tilde{p}(x_1) = \alpha_1$; autrement dit, \tilde{p} est un polynôme de degré 1 qui interpole α_0 au point x_0 et α_1 au point x_1 .

Définition 2.3.4. Le polynôme d'interpolation de Lagrange de degré 1, associé à f aux points x_0 et x_1 est

$$\tilde{p}(x) = \alpha_0 l_0(x) + \alpha_1 l_1(x) .$$

Proposition 2.3.5. Le polynôme de degré 1 qui interpole les valeurs de f aux points x_0 et x_1 est unique. En particulier les polynômes de Newton et de Lagrange sont deux expressions du même polynôme.

DÉMONSTRATION : En effet, supposons qu'il en existe un autre, disons q_1 . Alors

$$q_1(x_0) - p_1(x_0) = 0 \quad \text{et} \quad q_1(x_1) - p_1(x_1) = 0 .$$

Mais la différence $q_1 - p_1$ est aussi un polynôme de degré 1, disons de la forme $\alpha x + \beta$, et les formules ci-dessus veulent dire que ce polynôme a deux zéros. Ceci entraîne que $\alpha = \beta = 0$ et donc $q_1 = p_1$. ■

Il est indispensable de savoir évaluer l'erreur que l'on fait quand on remplace la fonction f par son polynôme d'interpolation de degré 1, p_1 .

Définition 2.3.6. L'erreur d'interpolation est définie par :

$$\text{Max}_{x \in I} |f(x) - p_1(x)|$$

où I est l'intervalle où on a choisi de remplacer f par p_1 .

Ici, on prendra un intervalle I qui contient les points d'interpolation x_0 et x_1 , mais ce n'est pas obligatoire.

Il est beaucoup plus difficile d'estimer l'erreur d'interpolation que d'estimer l'erreur d'approximation d'une dérivée en un point, comme on l'a fait au paragraphe précédent. La raison est évidemment que l'erreur d'interpolation demande une estimation en tous les points de l'intervalle I et non pas en un point seulement.

Proposition 2.3.7. Si f est deux fois dérivable, sur un intervalle de longueur h , l'erreur entre f et son polynôme d'interpolation p_1 qui prend les mêmes valeurs que f aux points x_0 et $x_0 + h$ est majorée par :

$$|f(x) - p_1(x)| \leq \frac{1}{2}h^2 M_2$$

où M_2 est tel que $\forall x \in [x_0, x_0 + h]$, $|f''(x)| \leq M_2$.

DÉMONSTRATION : Supposons que f est deux fois dérivable dans l'intervalle I et posons

$$g(x) = f(x) - p_1(x)$$

Cette fonction est aussi deux fois dérivable. Plaçons nous dans le cas le point x parcourt l'intervalle $[x_0, x_1]$. Comme g s'annule aux deux extrémités de cet intervalle, il est facile de remarquer que g ou $-g$ ne peut atteindre son maximum sur cet intervalle qu'en un point où la fonction g' s'annule.

Soit α un point où g' s'annule et appliquons la Formule de Taylor d'ordre $n = 2$ à g avec $a = \alpha$ et $b = x_0$:

$$g(x_0) = g(\alpha) + (x_0 - \alpha)g'(\alpha) + \frac{1}{2}(x_0 - \alpha)^2 g''(c_0),$$

où c_0 est un point compris entre x_0 et α et donc forcément $x_0 < c_0 < \alpha$. Mais comme $g(x_0) = 0$ et $g'(\alpha) = 0$, cette égalité nous donne

$$g(\alpha) = -\frac{1}{2}(x_0 - \alpha)^2 g''(c_0).$$

On peut se demander pourquoi on a privilégié x_0 et on peut souhaiter faire le même calcul avec $b = x_1$. On trouve

$$g(x_1) = g(\alpha) + (x_1 - \alpha)g'(\alpha) + \frac{1}{2}(x_1 - \alpha)^2 g''(c_1),$$

où c_1 est un point compris entre α et x_1 , c'est-à-dire $\alpha < c_1 < x_1$. Donc, on a aussi

$$g(\alpha) = -\frac{1}{2}(x_1 - \alpha)^2 g''(c_1).$$

Mais $g = f - p_1$ et $p_1'' = 0$, puisque c'est un polynôme de degré 1 ; donc ces équations deviennent

$$f(\alpha) - p_1(\alpha) = -\frac{1}{2}(x_0 - \alpha)^2 f''(c_0)$$

$$f(\alpha) - p_1(\alpha) = -\frac{1}{2}(x_1 - \alpha)^2 f''(c_1)$$

et ces égalités sont valables pour n'importe quel point α entre x_0 et x_1 où $(f - p_1)'(\alpha) = 0$, donc en particulier où $f - p_1$ ou $p_1 - f$ atteignent leur maximum.

Evidemment, si la distance entre x_0 et x_1 est grande, ces formules ne disent pas que p_1 est proche de f . Mais lorsque la distance entre x_0 et x_1 est petite, par exemple lorsque $x_1 = x_0 + h$, où h est destiné à tendre vers zéro, alors on en déduit la majoration

$$|f(x) - p_1(x)| \leq \frac{1}{2}h^2 M_2$$

pourvu que

$$\forall x \in [x_0, x_0 + h], |f''(x)| \leq M_2.$$

■

Cette démonstration, assez naturelle, est particulière au degré 1. Il existe une démonstration générale, mais beaucoup moins naturelle, qui s'adapte à des polynômes de degré quelconque, que nous verrons dans la suite de ce chapitre, au paragraphe 2.5.

Dans le cas du degré 1, cette démonstration générale montre que pour chaque point x dans un intervalle I contenant x_0 et x_1 , il existe un point c dans cet intervalle où on a l'égalité

$$f(x) - p_1(x) = \frac{1}{2}(x - x_0)(x - x_1)f''(c)$$

C'est un résultat plus fort que les inégalités précédentes, puisqu'il concerne n'importe quel point de I et pas seulement les points où $f' = p_1'$.

De plus, dans le cas simplifié où $x \in [x_0, x_1]$,

$$\sup_{x \in [x_0, x_1]} |(x - x_0)(x - x_1)| = \sup_{x \in [x_0, x_1]} (x - x_0)(x_1 - x) = \sup_{x \in [x_0, x_1]} -x^2 + (x_0 + x_1)x - x_0x_1.$$

Une étude facile de cette fonction montre que ce maximum est atteint au milieu de l'intervalle, c'est à dire en $x_m = (x_0 + x_1)/2$ et alors

$$\sup_{x \in [x_0, x_1]} |(x - x_0)(x - x_1)| = \frac{1}{4}(x_1 - x_0)^2.$$

Donc, en écrivant $x_1 = x_0 + h$, et en majorant la relation

$$f(x) - p_1(x) = \frac{1}{2}(x - x_0)(x - x_1)f''(c)$$

on trouve :

$$\forall x \in [x_0, x_0 + h], |f(x) - p_1(x)| \leq \frac{1}{8}h^2 M_2,$$

ce qui est une meilleure estimation que 2.3.7.

Extension. On cherche à approcher une courbe par une ligne polygonale : on suppose que cette courbe est le graphe d'une fonction f définie sur un intervalle $[a, b]$.

On découpe l'intervalle $[a, b]$ en sous intervalles $[x_i, x_{i+1}]$, avec $x_0 = a$, $x_1 = a + h$, ..., $x_n = b = a + nh$, où $h > 0$ est donné. Sur chaque sous intervalle $[x_i, x_{i+1}]$, on remplace f par son polynôme d'interpolation de degré 1 qui prend les mêmes valeurs que f aux points x_i et x_{i+1} .

La fonction Π_f ainsi construite sur $[a, b]$ est continue et c'est un polynôme par morceaux. Si on se reporte à la majoration 2.3.7, on voit que si

$$\forall x \in I, |f''(x)| \leq M_2,$$

alors

$$\forall x \in I, |f(x) - \Pi_f(x)| \leq \frac{1}{2}h^2 M_2.$$

Π_f est donc une approximation de f avec une erreur en $O(h^2)$ où h est la longueur de chaque sous-intervalle. Si on veut que cette approximation soit convenablement précise, il faut donc que chaque sous-intervalle soit assez petit, c'est à dire que h soit petit.

2.4 Interpolation d'une fonction par un polynôme de degré 2.

La construction directe d'un polynôme qui prend les mêmes valeurs qu'une fonction f , en trois points distincts deux à deux x_0 , x_1 et x_2 , n'est pas aussi facile que la construction d'une droite qui passe par 2 points. Mais nous avons vu que le polynôme d'interpolation de degré 1 a plusieurs expressions et nous pouvons nous en inspirer. Des deux expressions que nous avons trouvées pour ce polynôme, c'est la formule de Lagrange 2.3.4 qui est la plus facile à généraliser. En effet elle s'écrit avec les polynômes de base l_0 et l_1 qui ont la propriété de s'annuler en l'un des points et de valoir 1 en l'autre point. Or cette propriété s'étend facilement à un ensemble de plusieurs points distincts. En particulier, dans le cas de 3 points, on cherche trois polynômes de base l_0 , l_1 et l_2 , de degré 2, qui vérifient

$$\begin{aligned}l_0(x_0) &= 1, \quad l_0(x_1) = 0, \quad l_0(x_2) = 0, \\l_1(x_0) &= 0, \quad l_1(x_1) = 1, \quad l_1(x_2) = 0, \\l_2(x_0) &= 0, \quad l_2(x_1) = 0, \quad l_2(x_2) = 1.\end{aligned}$$

Considérons le cas de l_0 ; c'est un polynôme de degré 2 qui a les deux racines x_1 et x_2 . Donc il a les deux facteurs $(x - x_1)$ et $(x - x_2)$ et aucun autre facteur puisque son degré est 2 :

$$l_0(x) = \alpha(x - x_1)(x - x_2).$$

La constante α est déterminée par le fait que $l_0(x_0) = 1$:

$$1 = l_0(x_0) = \alpha(x_0 - x_1)(x_0 - x_2),$$

d'où

$$\alpha = \frac{1}{(x_0 - x_1)(x_0 - x_2)},$$

et

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}$$

En employant la même démarche pour calculer l_1 et l_2 , on trouve

$$\begin{aligned}l_1(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} \\l_2(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}\end{aligned}$$

Définition 2.4.1. Les polynômes l_0, l_1, l_2 sont appelés les polynômes de base de degré 2 associés à x_0, x_1, x_2 .

Notons que le graphe dans \mathbb{R}^2 du polynôme de degré 2

$$p_2(x) = f(x_0)l_0(x) + f(x_1)l_1(x) + f(x_2)l_2(x)$$

passé par les points $(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$.

Définition 2.4.2. *Le polynôme d'interpolation de Lagrange de degré 2 associé à f est :*

$$p_2(x) = f(x_0)l_0(x) + f(x_1)l_1(x) + f(x_2)l_2(x).$$

Comme précédemment, on a l'unicité du polynôme d'interpolation de degré 2 :

Proposition 2.4.3. *Il n'existe pas d'autre polynôme de degré 2 prenant les valeurs $f(x_i)$, $i = 0, 1, 2$ aux points x_0, x_1, x_2 .*

DÉMONSTRATION : En effet, s'il en existe un autre, disons q_2 , la différence $q_2 - p_2$ s'annule aux trois points distincts x_0, x_1 et x_2 . Comme son degré est 2, ce polynôme est nul et donc $q_2 = p_2$. ■

Pour écrire le polynôme d'interpolation de Lagrange sous la forme de Newton, analogue à l'ordre 1, il faut introduire les différences divisées d'ordre deux. La formule 2.3.1 définit la différence divisée d'ordre un de f aux points x_0 et x_1 . De même, la différence divisée d'ordre 1 de f aux points x_1 et x_2 est

$$f[x_1, x_2] = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Définition 2.4.4. *La différence divisée d'ordre 2 de f aux points x_0, x_1 et x_2 est définie par :*

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}.$$

Maintenant, remarquons que

$$f[x_0, x_1](x - x_0)|_{x=x_1} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_1 - x_0) = f(x_1) - f(x_0),$$

et

$$f[x_0, x_1, x_2](x - x_0)|_{x=x_2} = f[x_1, x_2] - f[x_0, x_1].$$

En comparant avec 2.3.2, on est amené à considérer le polynôme de degré 2 :

$$\tilde{p}_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1).$$

On trouve immédiatement que

$$\begin{aligned} \tilde{p}_2(x_0) &= f(x_0), \\ \tilde{p}_2(x_1) &= f(x_0) + f(x_1) - f(x_0) = f(x_1). \end{aligned}$$

Puis

$$\tilde{p}_2(x_2) = f(x_0) + f[x_0, x_1](x_2 - x_0) + (f[x_1, x_2] - f[x_0, x_1])(x_2 - x_1) = f(x_2).$$

Donc \tilde{p}_2 est un polynôme de degré 2 qui passe par les 3 points $(x_0, f(x_0))$, $(x_1, f(x_1))$ et $(x_2, f(x_2))$. Comme il n'y a qu'un seul polynôme de degré 2 qui passe par ces points, c'est le polynôme p_2 .

Définition 2.4.5. *Le polynôme d'interpolation de Newton associé à la fonction f est*

$$p_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1).$$

L'erreur d'interpolation de f par le polynôme p_2 , c'est-à-dire

$$\text{Max}_{x \in I} |f(x) - p_2(x)|,$$

où I est un intervalle, est difficile à estimer. Le résultat général que nous démontrerons au paragraphe suivant (Proposition 2.5.9) permet de prouver que, si f est 3 fois dérivable, pour chaque x dans un intervalle I contenant les points x_0 , x_1 et x_2 , il existe un point c dans cet intervalle où on a l'égalité

$$f(x) - p_2(x) = \frac{1}{6}(x - x_0)(x - x_1)(x - x_2)f'''(c)$$

En admettant provisoirement ce résultat, on obtient :

Proposition 2.4.6. *Si f est 3 fois dérivable, l'erreur d'interpolation entre f et son polynôme d'interpolation de degré 2 est majorée par :*

$$\text{Max}_{x \in [x_0, x_0 + 2h]} |f(x) - p_2(x)| < \frac{1}{3}h^3 M_3,$$

où M_3 est une borne supérieure de f''' dans l'intervalle $[x_0, x_0 + 2h]$.

DÉMONSTRATION : En effet, pour simplifier, on suppose que $x_1 = x_0 + h$ et $x_2 = x_1 + h = x_0 + 2h$ et que $x_0 < x < x_2$. Alors le produit $|(x - x_0)(x - x_1)(x - x_2)|$ est certainement inférieur à $2h^3$. C'est une majoration grossière, que l'on pourrait raffiner (en fait, sa borne supérieure est $\frac{2\sqrt{3}}{9}h^3$), mais néanmoins elle nous donne bien

$$\text{Max}_{x \in [x_0, x_0 + 2h]} |f(x) - p_2(x)| < \frac{1}{3}h^3 M_3,$$

où M_3 est une borne supérieure de f''' dans l'intervalle $[x_0, x_0 + 2h]$. ■

Il est souvent plus commode de calculer numériquement le polynôme d'interpolation sous la forme de Newton que sous la forme de Lagrange. Il est donc important de savoir calculer efficacement les différences divisées et de comprendre leurs propriétés.

Remarquons d'une part que dans la définition 2.4.4, l'ordre des points x_0 , x_1 et x_2 est sans importance, c'est-à-dire

$$f[x_0, x_1, x_2] = f[x_0, x_2, x_1] = f[x_1, x_0, x_2] = \dots = f[x_2, x_1, x_0].$$

D'autre part, cherchons une relation entre $f[x_0, x_1, x_2]$ et les différences finies. Pour cela, on suppose que $x_1 = x_0 + h$, $x_2 = x_1 + h = x_0 + 2h$. Alors

$$f[x_0, x_1, x_2] = \frac{1}{2h} \left(\frac{\Delta f(x_1)}{h} - \frac{\Delta f(x_0)}{h} \right) = \frac{1}{2} \frac{1}{h^2} \Delta^2 f(x_0).$$

Avec les techniques du paragraphe précédent, on peut montrer que

$$\lim_{h \rightarrow 0} \frac{1}{h^2} \Delta^2 f(x_0) = f''(x_0).$$

Donc

$$\lim_{h \rightarrow 0} f[x_0, x_0 + h, x_0 + 2h] = \frac{1}{2} f''(x_0).$$

En rapprochant cette limite de la forme 2.4.5, on voit que la forme de Newton ressemble un peu à un développement de Taylor.

Enfin, pour organiser efficacement le calcul de la différence divisée d'ordre deux de f aux points x_0, x_1 et x_2 , on peut disposer les nombres selon le tableau suivant :

$$\begin{array}{cccc} x_0 & f[x_0] & & \\ x_1 & f[x_1] & f[x_0, x_1] & \\ x_2 & f[x_2] & f[x_1, x_2] & f[x_0, x_1, x_2] \end{array}$$

Les deux premières colonnes sont données. A partir de ces deux colonnes, on calcule les éléments de la troisième colonne et à partir de la première et de la troisième colonne, on calcule les éléments (ici, il n'y a qu'un élément) de la quatrième colonne.

Dans le paragraphe suivant, on généralisera la définition des différences divisées à l'ordre n : elles servent à construire les polynômes d'interpolation de degré n . L'avantage de cette disposition des calculs est qu'elle s'adapte facilement au calcul des différences divisées d'ordre n et donc aux calculs des coefficients des polynômes d'interpolation passant par $n + 1$ points.

2.5 Interpolation d'une fonction par un polynôme de degré n

La construction directe d'un polynôme qui prend les mêmes valeurs qu'une fonction f , en $n + 1$ points distincts deux à deux $x_0, x_1 \dots x_n$ est, comme dans le cas $n = 2$, plus facile à partir de la formule de Lagrange 2.3.4 et 2.4.2.

On généralise les polynômes de base $l_0, l_1 \dots l_n$ pour qu'ils soient de degré n qu'ils aient la propriété de s'annuler en tous les points sauf 1 où ils valent 1. On pose :

$$l_i(x) = \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)}, \quad 0 \leq i \leq n$$

On vérifie très facilement que $l_i \in \mathcal{P}_n$ pour tout $0 \leq i \leq n$ et que $l_i(x_i) = 1, l_i(x_j) = 0, j \neq i$.

Définition 2.5.1. Les polynômes l_0, l_1, \dots, l_n sont appelés les polynômes de base de degré n associés à x_0, x_1, \dots, x_n .

Définition 2.5.2. Le polynôme d'interpolation de Lagrange de degré n associé à f et aux $n + 1$ points x_0, x_1, \dots, x_n est :

$$p_n(x) = \sum_{i=0}^n f(x_i) l_i(x).$$

Comme précédemment, on a l'unicité de ce polynôme :

Proposition 2.5.3. *Il n'existe pas d'autre polynôme de degré n dont le graphe passe par les $n + 1$ points $(x_0, f(x_0)), (x_1, f(x_1)), \dots (x_n, f(x_n))$.*

DÉMONSTRATION : En effet, s'il en existe un autre, disons q_n , la différence $q_n - p_n$ s'annule aux $n + 1$ points distincts $x_0, x_1 \dots x_n$. Comme son degré est n , ce polynôme est nul et donc $q_n = p_n$. En conclusion, le polynôme d'interpolation de degré n passant par $n + 1$ points distincts deux à deux est unique. ■

Notations 2.5.4. *Dans toute la suite, on pose :*

$$\pi_{n+1}(x) = \prod_{j=0}^n (x - x_j).$$

On remarque qu'alors :

$$\text{pour } x \neq x_i, \quad l_i(x) = \frac{\pi_{n+1}(x)}{(x - x_i)\pi'_{n+1}(x_i)}$$

Pour écrire le polynôme d'interpolation de Lagrange sous la forme de Newton, d'une façon analogue aux ordres 1 et 2, il faut introduire les différences divisées d'ordre k , $k \leq n$. Les formules 2.3.1 et 2.4.4 définissent les différences divisées d'ordre 1 et 2 de f . On définit par récurrence, la différence divisée d'ordre k de f aux points $x_i, x_{i+1} \dots x_{i+k}$ par :

Définition 2.5.5. *La différence divisée d'ordre k , $k \leq n$ de f aux $k + 1$ points $x_i, x_{i+1} \dots x_{i+k}$, $0 \leq i \leq n - k$ est définie par :*

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}.$$

En comparant avec 2.3.2, on est amené à considérer :

Définition 2.5.6. *Le polynôme d'interpolation de Newton de degré n associé à la fonction f et aux $n + 1$ points x_0, x_1, \dots, x_n est*

$$q_n(x) = f[x_0] + \sum_{k=1}^n f[x_0, x_1, \dots, x_k](x - x_0)\dots(x - x_{k-1}).$$

Pour tout $k \leq n$, notons $g[x_0, \dots, x_k]$ le coefficient directeur de x^k dans le polynôme de Lagrange p_k de f . Comme le polynôme $p_k - p_{k-1}$ est de degré k et s'annule aux k points x_0, \dots, x_{k-1} , on trouve immédiatement que

$$p_k(x) - p_{k-1}(x) = g[x_0, x_1, \dots, x_k](x - x_0)\dots(x - x_{k-1})$$

Comme $p_0(x_0) = f(x_0)$, on en déduit par récurrence que

$$p_n(x) = f[x_0] + \sum_{k=1}^n g[x_0, x_1, \dots, x_k](x - x_0)\dots(x - x_{k-1}).$$

Pour montrer que le polynôme de Newton de f coïncide avec celui de Lagrange, il suffit donc de montrer :

Proposition 2.5.7. Pour tout $k \leq n$, le coefficient directeur de x^k dans le polynôme de Lagrange p_k est égal à $f[x_0, x_1, \dots, x_k]$.

DÉMONSTRATION : On procède par récurrence sur k . Désignons par $r_{k-1} \in \mathcal{P}_{k-1}$ le polynôme d'interpolation de f aux points : x_1, x_2, \dots, x_k . Posons :

$$\tilde{p}_k(x) = \frac{(x - x_0)r_{k-1}(x) - (x - x_k)p_{k-1}(x)}{x_k - x_0}$$

Alors, $\tilde{p}_k \in \mathcal{P}_k$, $\tilde{p}_k(x_0) = p_{k-1}(x_0) = f(x_0)$, $\tilde{p}_k(x_k) = r_{k-1}(x_k) = f(x_k)$ et pour tout $0 < i < k$, on a :

$$\tilde{p}_k(x_i) = \frac{(x_i - x_0)f(x_i) - (x_i - x_k)f(x_i)}{x_k - x_0} = f(x_i)$$

Par conséquent, $\tilde{p}_k = p_k$. Comme le coefficient directeur de r_{k-1} est $f[x_1, x_2, \dots, x_k]$, on obtient le résultat cherché en égalant les coefficients de x^k dans l'identité :

$$p_k(x) = \frac{(x - x_0)r_{k-1}(x) - (x - x_k)p_{k-1}(x)}{x_k - x_0}$$

■

Corollaire 2.5.8. Le polynôme de Newton d'ordre n de f , associé aux $n + 1$ points x_0, x_1, \dots, x_n coïncide avec son polynôme d'interpolation de Lagrange d'ordre n .

Nous allons maintenant évaluer l'erreur d'interpolation de f par le polynôme p_n , c'est-à-dire

$$\text{Max}_{x \in I} |f(x) - p_n(x)|,$$

où I est un intervalle sur lequel p_n remplace f .

Proposition 2.5.9. Si f est $n + 1$ fois dérivable sur $[a, b]$, l'erreur d'interpolation entre f et son polynôme d'interpolation de degré n est égale à :

$$f(x) - p_n(x) = \frac{1}{(n + 1)!} \pi_{n+1}(x) f^{(n+1)}(\xi_x),$$

où ξ_x est un point de l'intervalle $]\min_{i=0, \dots, n}(x, x_i), \max_{i=0, \dots, n}(x, x_i)[$.

DÉMONSTRATION : On a besoin d'un lemme :

Lemme 2.5.10. Soit g une fonction p fois dérivable sur $[a, b]$. On suppose qu'il existe $p + 1$ points $c_0 < c_1 < \dots < c_p$ de $[a, b]$ tels que $g(c_i) = 0$. Alors il existe $\xi \in]c_0, c_p[$ tel que $g^{(p)}(\xi) = 0$.

DÉMONSTRATION : La démonstration se fait par récurrence sur p . Pour $p = 1$, c'est le théorème de Rolle. Supposons que le résultat est vrai à l'ordre $p - 1$. Sous l'hypothèse du Lemme 2.5.10, le théorème de Rolle implique l'existence de points $\gamma_0 \in]c_0, c_1[$, ..., $\gamma_{p-1} \in]c_{p-1}, c_p[$ tels que $g'(\gamma_i) = 0$. Par l'hypothèse de récurrence, il existe donc $\xi \in]\gamma_0, \gamma_p[$ tel que $(g')^{(p-1)}(\xi) = g^{(p)}(\xi) = 0$. ■

Revenons à la démonstration de la Proposition 2.5.9 :

-Si $x = x_i$, on a $\pi_{n+1}(x_i) = 0$ et tout point ξ_x convient.

-Supposons $x \neq x_i$. Considérons la fonction

$$\forall t \in [a, b], g(t) = f(t) - p_n(t) - c\pi_{n+1}(t)$$

où c est choisi tel que $g(x) = 0$, c'est à dire $c = \frac{f(x) - p_n(x)}{\pi_{n+1}(x)}$.

Cette fonction s'annule en les $n + 2$ points x, x_0, x_1, \dots, x_n donc d'après le Lemme 2.5.10, il existe $\xi_x \in]\min(x, x_i), \max(x, x_i)[$ tel que $g^{(n+1)}(\xi_x) = 0$. Or $p_n \in \mathcal{P}_n$ donc $p_n^{(n+1)} = 0$ et il est facile de voir que $\pi_{n+1}^{(n+1)} = (n + 1)!$. On en déduit :

$$g^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - c(n + 1)! = 0 \Rightarrow c = \frac{f^{(n+1)}(\xi_x)}{(n + 1)!}$$

D'où

$$f(x) - p_n(x) = p_{n+1}(x) - p_n(x) = c\pi_{n+1}(x) = \frac{f^{(n+1)}(\xi_x)}{(n + 1)!}\pi_{n+1}(x)$$

■

Si on connaît une borne pour les dérivées de f et les polynômes π_n , ceci donne une majoration de l'erreur, précisément :

$$\|f - p_n\| \leq \frac{1}{(n + 1)!} \|\pi_{n+1}\| \|f^{(n+1)}\|$$

Si les points x_i sont quelconques, seule la majoration triviale pour $\|\pi_{n+1}\|$ est possible c'est à dire

$$\|\pi_{n+1}\| \leq (b - a)^{n+1}$$

ce qui donne :

Corollaire 2.5.11. *Quelques soient les points $x_i, i = 0 \dots n$, si la fonction f est de classe $C^{(n+1)}$, on a*

$$\|f - p_n\| \leq \frac{1}{(n + 1)!} \|\pi_{n+1}\| \|f^{(n+1)}\| \leq \frac{(b - a)^{n+1}}{(n + 1)!} \|f^{(n+1)}\|$$

Enfin, pour organiser efficacement le calcul des différences divisées d'ordre n de f aux points $x_0, x_1 \dots x_n$, on peut disposer les nombres selon le tableau suivant :

[0]	$f[x_0]$				
[1]	$f[x_1]$	$f[x_0, x_1]$			
[2]	$f[x_2]$	$f[x_1, x_2]$	$f[x_0, x_1, x_2]$		
...					
[n - 2]	$f[x_{n-2}]$				
[n - 1]	$f[x_{n-1}]$	$f[x_{n-2}, x_{n-1}]$			
[n]	$f[x_n]$	$f[x_{n-1}, x_n]$	$f[x_{n-2}, x_{n-1}, x_n]$...	$f[x_0, \dots, x_n]$

On calcule les éléments d'une colonne à partir de la précédente par la formule 2.5.5.

Corollaire 2.5.12. Si l'on suppose que $f, f', f'', \dots, f^{(n)}$ existent et sont continues, les différences divisées $f[x_i, x_{i+1}, \dots, x_{i+k}]$ pour $0 \leq i \leq i+k \leq n$, sont bornées indépendamment du choix des x_j pour $i \leq j \leq i+k$. Précisément, il existe $\xi \in [x_i, x_{i+k}]$ tel que :

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{1}{k!} f^{(k)}(\xi)$$

DÉMONSTRATION : Pour simplifier la présentation, on suppose que $i = 0$, le cas général s'en déduisant sans peine en changeant l'indexation des points.

On a vu dans 2.5.7 que :

$$p_k(x) - p_{k-1}(x) = f[x_0, x_1, \dots, x_k](x - x_0) \dots (x - x_{k-1})$$

Donc au point x_k :

$$p_k(x_k) - p_{k-1}(x_k) = f[x_0, x_1, \dots, x_k](x_k - x_0) \dots (x_k - x_{k-1})$$

Or la formule d'erreur appliquée en x_k à l'ordre $k - 1$ donne :

$$p_k(x_k) - p_{k-1}(x_k) = f(x_k) - p_{k-1}(x_k) = \frac{1}{k!} \pi_k(x_k) f^{(k)}(\xi),$$

où ξ est un point de l'intervalle $[\min(x_0, \dots, x_k), \max(x_0, \dots, x_k)]$.

En comparant les deux égalités, on voit que :

$$f[x_0, x_1, \dots, x_k] = \frac{1}{k!} f^{(k)}(\xi),$$

ce qui implique le résultat. ■

Cas Particulier 2.5.13. Comme dans les cas d'interpolation d'ordre 1 et 2, on considère le cas où les points d'interpolation sont équidistants, c'est à dire que l'on prend un pas constant $h = \frac{b-a}{n}$ et que l'on définit :

$$x_i = a + ih = a + i \frac{b-a}{n}, \quad 0 \leq i \leq n$$

On note $f_i = f(x_i)$ les valeurs de f correspondantes et on détermine, comme au paragraphe précédent, les relations entre les différences finies et les différences divisées :

Définition 2.5.14. On définit l'opérateur aux différences finies Δ de (f_0, f_1, \dots, f_n) dans $(\Delta f_0, \Delta f_1, \dots, \Delta f_{n-1})$ par : $\Delta f_i = f_{i+1} - f_i$.

Par récurrence, les itérés de Δ vérifient :

$$\Delta^k f_i = \Delta^{k-1} f_{i+1} - \Delta^{k-1} f_i$$

Proposition 2.5.15.

- 1) On a : $\Delta^k f_i = \sum_{j=0}^k (-1)^{k-j} C_k^j f_{i+j}$.
- 2) $f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{\Delta^k f_i}{k!h^k}$
- 3) Si f est de classe C^k , on a : $\lim_{h \rightarrow 0} \frac{\Delta^k f_i}{k!h^k} = \frac{f^{(k)}(x_i)}{k!}$.

DÉMONSTRATION : 1) se vérifie sans problème et 2) se démontre par récurrence à partir de 1).

Pour démontrer 3), on utilise l'expression trouvée en 2) $f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{\Delta^k f_i}{k!h^k}$ et on applique le Corollaire 2.5.12 : il existe $\xi \in [x_i, x_{i+k}]$ tel que

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{\Delta^k f_i}{k!h^k} = \frac{1}{k!} f^{(k)}(\xi)$$

Lorsque $h \rightarrow 0$, $\xi \rightarrow x_i$ et par continuité,

$$\lim_{h \rightarrow 0} \frac{\Delta^k f_i}{k!h^k} = \frac{1}{k!} f^{(k)}(x_i)$$

■

En ré-écrivant le polynôme d'interpolation sous forme de Newton, Définition 2.5.6, on obtient :

Proposition 2.5.16. On pose $x = x_0 + sh$, $s \in [0, n]$, alors :

$$\begin{aligned} p_n(x) &= \sum_{k=0}^n \Delta^k f_0 \frac{s(s-1)\dots(s-k+1)}{k!} \\ &= f_0 + \frac{s}{1} (\Delta^1 f_0 + \frac{s-1}{2} (\Delta^2 f_0 + \dots + \frac{s-n+1}{n} (\Delta^n f_0) \dots)). \end{aligned}$$

DÉMONSTRATION : Il suffit d'écrire, avec les notations précédentes :

$$(x - x_0)(x - x_1)\dots(x - x_{k-1}) = h^k s(s-1)(s-2)\dots(s-k+1)$$

et appliquer la formule 2) de la Proposition 2.5.14.

■

Les coefficients $\Delta^k f_0$ se calculent selon le schéma des différences finies, que l'on peut représenter par le tableau :

$$\begin{array}{ccccccc} f_0 & & & & & & \\ f_1 & \Delta f_0 & & & & & \\ f_2 & \Delta f_1 & \Delta^2 f_0 & & & & \\ \dots & & & & & & \\ f_{n-2} & & & & & & \\ f_{n-1} & \Delta f_{n-2} & & & & & \\ f_n & \Delta f_{n-1} & \Delta^2 f_{n-2} & \dots & \Delta^n f_0 & & \end{array}$$

On va maintenant estimer l'erreur d'interpolation dans le cas des points équidistants : Etant donné que $\pi_{n+1}(x) = h^{n+1}s(s-1)(s-2)\dots(s-n)$, on a, en utilisant 2.5.9

$$f(x) - p_n(x) = \frac{s(s-1)(s-2)\dots(s-n)}{(n+1)!} h^{n+1} f^{(n+1)}(\xi_x)$$

La fonction $\phi(s) = |s(s-1)(s-2)\dots(s-n)|$ pour $s \in [0, n]$ vérifie : $\phi(n-s) = \phi(s)$, elle atteint donc son maximum dans $[0, n/2]$. Comme de plus, $\frac{\phi(s-1)}{\phi(s)} = \frac{n+1-s}{s} > 1$ pour $1 \leq s \leq \frac{n}{2}$, on voit que ϕ atteint son maximum sur $[0, 1]$, d'où :

$$\max\{\phi(s) \mid s \in [0, n]\} = \max\{\phi(s) \mid s \in [0, 1]\} \leq n!$$

Il en résulte que

$$|f(x) - p_n(x)| = \frac{1}{n+1} h^{n+1} \max\{f^{(n+1)}(\xi) \mid \xi \in [x_0, x_1, \dots, x_n]\}$$

Remarque 2.5.17. Ce calcul d'estimation de ϕ donne aussi une estimation de $\|\pi_{n+1}\|$ dans le cas de points équidistants :

$$\|\pi_{n+1}\| = h^{n+1} \max\{\phi(s) \mid s \in [0, n]\} \leq h^{n+1} n! = \frac{n!}{n^{n+1}} (b-a)^{n+1}$$

et par la formule de Stirling : $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, on voit que l'ordre de grandeur de $\|\pi_{n+1}\|$ est :

$$\|\pi_{n+1}\| \leq O\left(\frac{b-a}{e}\right)^{n+1}$$

quand $n \rightarrow +\infty$.

Cette estimation est meilleure que celle obtenue pour des points quelconques, cf. Corollaire 2.5.11.

On peut également utiliser d'autres points $x_i, i = 0, \dots, n$, par exemple les points de Tchebichev :

Cas Particulier 2.5.18. Interpolation aux points de Tchebichev : on définit les polynômes de Tchebichev par : $t_n(u) = \cos(n \arccos u)$, $u \in [-1, +1]$.

Montrons d'abord que t_n est un polynôme : on pose $\theta = \arccos u$ soit $u = \cos \theta$ avec $\theta \in [0, \pi]$. On a alors : $t_n(u) = \cos n\theta$ et par suite :

$$t_{n+1}(u) + t_{n-1}(u) = \cos((n+1)\theta) + \cos((n-1)\theta) = 2 \cos n\theta \cos \theta = 2ut_n(u)$$

Les fonctions t_n se calculent donc par récurrence, avec les formules :

$$t_0(u) = 1, \quad t_1(u) = u, \quad t_{n+1}(u) = 2ut_n(u) - t_{n-1}(u)$$

Il en résulte que t_n est bien un polynôme de degré n pour tout n , dont le coefficient directeur est 2^{n-1} si $n \geq 1$.

Les racines de t_n dans $[-1, +1]$ sont déterminées par :

$$t_n(u) = \cos n\theta = 0 \iff n\theta = \frac{\pi}{2} + k\pi \iff \theta = \frac{2k+1}{2n}\pi, \quad 0 \leq k \leq n-1$$

Le polynôme t_n admet donc exactement n racines distinctes dans $[-1, +1]$ données par

$$u_k = \cos \frac{2k+1}{2n}\pi, \quad 0 \leq k \leq n-1$$

Comme t_n est de degré n , il ne peut pas avoir d'autres racines.

Définition 2.5.19. Les points d'interpolation de Tchebichev d'ordre n sont les $n+1$ points $u_k = \cos \frac{2k+1}{2(n+1)}\pi$, $0 \leq k \leq n$, racines du polynôme t_{n+1} .

Les points u_k vérifient : $u_{n-k} = u_k$ pour $k = 0, 1, \dots, n$, ils sont donc répartis symétriquement autour de 0.

De plus, puisque le coefficient directeur de t_{n+1} est 2^n , il vient :

$$t_{n+1}(u) = 2^n \prod_{k=0}^n (u - u_k) = 2^n \pi_{n+1}(u)$$

Pour se ramener à un intervalle quelconque $[a, b]$ au lieu de $[-1, +1]$, on utilisera la bijection affine de $[-1, +1]$ sur $[a, b]$ définie par $u \rightarrow x = \frac{a+b}{2} + \frac{b-a}{2}u$ qui envoie -1 sur a et $+1$ sur b . Les images des $n+1$ points d'interpolation de Tchebichev $u_k \in [-1, +1]$, $0 \leq k \leq n$ sont donnés par

$$x_k = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{2k+1}{2(n+1)}\pi, \quad 0 \leq k \leq n$$

Définition 2.5.20. Les $n+1$ points $x_k = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{2k+1}{2(n+1)}\pi$, $0 \leq k \leq n$, sont appelés points d'interpolation de Tchebichev d'ordre n de l'intervalle $[a, b]$.

Pour ces points, on a $x - x_k = \frac{b-a}{2}(u - u_k)$ et donc le polynôme π_{n+1} est donné par :

$$\begin{aligned} \pi_{n+1}(x) &= \prod_{k=0}^n (x - x_k) = \left(\frac{b-a}{2}\right)^{n+1} \prod_{i=0}^n (u - u_k) = \\ &= \left(\frac{b-a}{2}\right)^{n+1} \frac{1}{2^n} t_{n+1}(u) = \frac{(b-a)^{n+1}}{2^{2n+1}} t_{n+1}\left(\frac{2}{b-a}\left(x - \frac{a+b}{2}\right)\right) \end{aligned}$$

Par définition des polynômes de Tchebichev, on a : $\|t_{n+1}\| = 1$, donc

$$\|\pi_{n+1}\| = 2\left(\frac{b-a}{4}\right)^{n+1}$$

Remarque 2.5.21. L'estimation $\|\pi_{n+1}\| = 2\left(\frac{b-a}{4}\right)^{n+1}$ obtenue pour les points de Tchebichev est meilleure que celle obtenue pour des points équidistants (2.5.17)

On peut chercher à démontrer la convergence uniforme des polynômes d'interpolation d'une fonction f vers cette fonction. Evidemment, cela dépend du choix des points d'interpolation et de la régularité de la fonction. Donnons un cas de convergence dans le cas où f est analytique :

Théorème 2.5.22. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction analytique donnée par une série entière de rayon de convergence $R > \frac{b-a}{2}$, centrée au point $\frac{a+b}{2}$. Alors, pour des points d'interpolation d'ordre n , $\{x_{0,n}, x_{1,n}, \dots, x_{n,n}\}$, les polynômes d'interpolation p_n convergent

uniformément vers f sur $[a, b]$ pourvu que $R > (\frac{1}{\lambda} + \frac{1}{2})(b - a)$ où $\lambda = 1$ en général, $\lambda = e$ pour des points équidistants et $\lambda = 4$ pour des points de Tchebichev.

DÉMONSTRATION : Pour $x \in [a, b]$, on pose $u = x - \frac{a+b}{2}$ et on a $|u| \leq \frac{b-a}{2}$.

Supposons que $f(u) = \sum_{k=0}^{+\infty} a_k u^k$, avec un rayon de convergence $R > \frac{b-a}{2}$.

Soit $r > 0$ tel que $\frac{b-a}{2} < r < R$. La série numérique $\sum_{k=0}^{+\infty} a_k r^k$ converge, donc son terme général est borné (en fait, il tend vers 0 à l'infini), soit

$$|a_k| \leq \frac{C(r)}{r^k}, \quad \forall k \in \mathbb{N}.$$

On peut dériver terme à terme la série sur $[-r, +r]$, ce qui donne

$$f^{(n)}(u) = \sum_{k=0}^{+\infty} a_k \frac{d^n}{du^n} u^k$$

et si $u \geq 0$:

$$\begin{aligned} |f^{(n)}(u)| &\leq C(r) \frac{d^n}{du^n} \left(\sum_{k=0}^{+\infty} \left(\frac{u}{r}\right)^k \right) \\ &= C(r) \frac{d^n}{du^n} \left(\frac{1}{1 - \frac{u}{r}} \right) = C(r) \frac{d^n}{du^n} \left(\frac{r}{r-u} \right) \\ &= \frac{n! r C(r)}{(r-u)^{n+1}}. \end{aligned}$$

Si $u \leq 0$, on pose $v = -u$ et on dérive par rapport à v , ce qui donne la même estimation.

En posant, pour toute fonction continue g , définie et continue sur $[-\frac{b-a}{2}, \frac{b-a}{2}]$:

$$\|g\|_{\infty} = \sup\{|g(u)|, |u| \leq \frac{b-a}{2}\}$$

pour tout r tel que $\frac{b-a}{2} < r < R$ et pour tout $n \in \mathbb{N}$, d'après ce qui précède :

$$\frac{1}{n!} \|f^{(n)}\|_{\infty} \leq \frac{rC(r)}{(r - \frac{b-a}{2})^{n+1}}$$

et l'erreur d'interpolation admet la majoration suivante :

$$\|f - p_n\|_{\infty} \leq \frac{1}{(n+1)!} \|\pi_{n+1}\| \|f^{(n+1)}\|_{\infty} \leq 2 \left(\frac{b-a}{\lambda}\right)^{n+1} \frac{rC(r)}{(r - \frac{b-a}{2})^{n+2}}$$

$$\leq \frac{2rC(r)}{r - \frac{b-a}{2}} \left(\frac{\frac{b-a}{\lambda}}{r - \frac{b-a}{2}} \right)^{n+1}$$

avec respectivement $\lambda = 1$ si les points d'interpolation sont quelconques, $\lambda = e$ s'ils sont équidistants et $\lambda = 4$ si ce sont les points de Tchebichev.

La suite $(p_n)_{n \in \mathbb{N}}$ va donc converger vers f uniformément sur $[a, b]$ si $\|f - p_n\|_\infty \rightarrow 0$ quand $n \rightarrow \infty$, c'est à dire si l'on peut choisir r tel que $\left(\frac{b-a}{\lambda}\right) < r - \frac{b-a}{2}$ soit $r > \left(\frac{1}{\lambda} + \frac{1}{2}\right)(b-a)$ et ceci est possible dès que le rayon de convergence R vérifie lui-même cette minoration. ■

Exemple 2.5.23. (Phénomène de Runge) Les polynômes d'interpolation de la fonction analytique

$$f_\alpha(x) = \frac{1}{x^2 + \alpha^2}, \quad x \in [-1, 1]$$

où $\alpha > 0$ ne forment pas une suite convergente pour toute valeur de α .

D'après ce qui précède, en prenant par exemple des points équidistants, si le rayon de convergence α vérifie : $\alpha > 2\left(\frac{1}{2} + \frac{1}{e}\right)$, la suite des polynômes d'interpolation de f_α converge uniformément vers f_α . Par contre, on peut montrer que si α est suffisamment petit, cette suite ne converge pas.

2.6 Stabilité numérique du procédé d'interpolation de Lagrange

Définition 2.6.1. Soient $x_0, x_1, \dots, x_n \in [a, b]$, $n + 1$ points deux à deux distincts. On définit l'opérateur d'interpolation de Lagrange associés à ces points, l'opérateur de $\mathcal{C}([a, b])$ dans \mathcal{P}_n défini par :

$$L_n(f) = p_n$$

où p_n est le polynôme d'interpolation de Lagrange de f associé aux points x_0, x_1, \dots, x_n .

Proposition 2.6.2. L_n est un opérateur linéaire de norme

$$\|L_n\| = \sup\left\{ \left(\sum_{i=0}^n |l_i(x)| \right) \mid x \in [a, b] \right\}$$

appelée constante de Lebesgue associée aux points x_0, x_1, \dots, x_n .

DÉMONSTRATION : Si p_n et r_n sont les polynômes d'interpolation associées à f et g respectivement, aux mêmes points x_0, x_1, \dots, x_n , on a $p_n(x_i) = f(x_i)$ et $r_n(x_i) = g(x_i)$. Donc pour tous réels λ, μ , $\lambda p_n(x_i) + \mu r_n(x_i) = \lambda f(x_i) + \mu g(x_i)$, donc $\lambda p_n + \mu r_n$ est le polynôme d'interpolation associé à $\lambda f + \mu g$ et L_n est bien linéaire.

D'autre part, on peut écrire, en reprenant la définition du polynôme d'interpolation d'ordre n , 2.5.2 :

$$p_n(x) = L_n(f)(x) = \sum_{i=0}^n f(x_i) l_i(x).$$

D'où :

$$|p_n(x)| \leq \left(\sum_{i=0}^n |l_i(x)| \right) \|f\|$$

et donc

$$\|p_n\| \leq \sup \left\{ \left(\sum_{i=0}^n |l_i(x)| \right) \mid x \in [a, b] \right\} \|f\| = \Lambda_n \|f\|.$$

Donc la norme est bien inférieure à $\Lambda_n = \sup \left\{ \left(\sum_{i=0}^n |l_i(x)| \right) \mid x \in [a, b] \right\}$.

Pour montrer l'égalité, par continuité des l_i , il existe un point $\xi \in [a, b]$ tel que $\Lambda_n = \sum_{i=0}^n |l_i(\xi)|$. Il suffit de prendre une fonction f_0 telle que $f_0(x_i) = \text{sgn}(l_i(\xi))$, $\|f_0\| = 1$. Alors :

$$L_n(f_0)(\xi) = \sum_{i=0}^n |l_i(\xi)| = \Lambda_n.$$

■

Remarque 2.6.3. La constante Λ_n peut s'interpréter comme le facteur d'amplification de l'erreur dans le procédé d'interpolation de Lagrange.

En effet, dans la pratique, on ne connaît pas toujours exactement la fonction f que l'on cherche à interpoler : on ne dispose que d'une valeur approchée $\tilde{f} = f + g$, où g est un terme d'erreur. Au lieu de calculer $p_n = L_n(f)$, on va calculer $\tilde{p}_n = L_n(\tilde{f}) = L_n(f) + L_n(g) = p_n + L_n(g)$. Si g est l'erreur commise sur f , $L_n(g)$ sera l'erreur commise sur p_n et on peut donc estimer $\|L_n(g)\|$ en fonction de g par $\|L_n(g)\| \leq \Lambda_n \|g\|$.

Cette constante Λ_n mesure également la proximité de f à son polynôme d'interpolation de degré n en fonction de sa distance aux polynômes de degré n :

Théorème 2.6.4. Pour tout $f \in \mathcal{C}([a, b])$, on a :

$$\|f - L_n(f)\| \leq (1 + \Lambda_n) d(f, \mathcal{P}_n)$$

où $d(f, \mathcal{P}_n)$ désigne la distance de f à l'espace \mathcal{P}_n pour la norme uniforme, c'est à dire : $d(f, \mathcal{P}_n) = \inf \{ \|f - p\| \mid p \in \mathcal{P}_n \}$.

DÉMONSTRATION : On a besoin d'un lemme :

Lemme 2.6.5. Pour tout $f \in \mathcal{C}([a, b])$, il existe un polynôme de meilleure approximation de degré n de f , c'est à dire un polynôme $q_n \in \mathcal{P}_n$ tel que $d(f, \mathcal{P}_n) = \|f - q_n\|$.

DÉMONSTRATION : En approximant f par 0, on voit que $d(f, \mathcal{P}_n) \leq \|f\|$. L'ensemble des polynômes $p \in \mathcal{P}_n$ tels que $\|f - p\| \leq \|f\|$ est une partie fermée bornée $K \subset \mathcal{P}_n$, non vide puisque $0 \in K$. Comme \mathcal{P}_n est de dimension finie, K est une partie compacte, donc la fonction continue $f \rightarrow \|f - p\|$ atteint son minimum en un point $q_n \in \mathcal{P}_n$. ■

Remarque. On peut montrer que ce polynôme de meilleure approximation est unique.

Revenons à la démonstration du théorème : Avec les notations précédentes, puisque q_n est un polynôme de degré n , on a $L_n(q_n) = q_n$, donc :

$$f - L_n(f) = f - q_n - L_n(f - q_n)$$

D'où

$$\begin{aligned} \|f - L_n(f)\| &\leq \|f - q_n\| + \|L_n(f - q_n)\| \\ &\leq \|f - q_n\| + \Lambda_n \|f - q_n\| = (1 + \Lambda_n)d(f, \mathcal{P}_n) \end{aligned}$$

■

Cas où les points sont équidistants. $x_i = a + ih$, $0 \leq i \leq n$ et $h = \frac{(b-a)}{n}$. On a alors :

$$\frac{2^n}{4n^2} \leq \Lambda_n \leq 2^n$$

DÉMONSTRATION : Comme en 2.5.13, on pose $x = a + sh$ où $s \in [0, n]$ et alors :

$$l_i(x) = \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)} = \prod_{j \neq i} \frac{(s - j)}{(i - j)} = (-1)^{n-i} \frac{s(s-1)\dots(\widehat{s-i})\dots(s-n)}{i!(n-i)!}$$

où $(\widehat{s-i})$ désigne un facteur omis.

Minoration : Pour $s = \frac{1}{2}$, il vient

$$\left| l_i\left(a + \frac{h}{2}\right) \right| = \frac{\frac{1}{2} \frac{1}{2} \frac{3}{2} \dots (\widehat{i - \frac{1}{2}}) \dots (n - \frac{1}{2})}{i!(n-i)!} \geq \frac{1}{4} \frac{1 \cdot 2 \dots (\widehat{i-1}) \dots (n-1)}{i!(n-i)!} \geq \frac{1}{4n^2} \frac{n!}{i!(n-i)!}$$

On en déduit :

$$\Lambda_n \geq \sum_{i=0}^n \left| l_i\left(a + \frac{h}{2}\right) \right| \geq \frac{1}{4n^2} \sum_{i=0}^n C_n^i = \frac{1}{4n^2} 2^n$$

Majoration : Pour l'autre inégalité, on écrit, pour $s \in [k, k+1]$, $0 \leq k \leq n$:

$$|l_i(a + sh)| \leq \frac{(n-k)!(k+1)!}{i!(n-i)!} \leq \frac{n!}{i!(n-i)!}$$

D'où

$$\Lambda_n = \sup \left\{ \sum_{i=0}^n |l_i(a + sh)| \mid s \in [k, k+1], 0 \leq k \leq n \right\} \leq \sum_{i=0}^n C_n^i = 2^n$$

■

Cas des points de Tchebichev. Dans ce cas, on a une meilleure estimation :

$$\frac{2}{\pi} \ln(n) \leq \Lambda_n \leq C \ln(n)$$

DÉMONSTRATION : Posons comme en 2.5.18, pour $u \in [-1, +1]$ et $\theta \in [0, \pi]$:

$$u = \cos \theta, \quad u_k = \cos \theta_k \text{ avec } \theta_k = \frac{2k+1}{2(n+1)}\pi, \quad 0 \leq k \leq n$$

La relation $t_{n+1}(\cos \theta) = \cos(n+1)\theta$ entraîne par dérivation

$$\sin \theta t'_{n+1}(\cos \theta) = (n+1) \sin(n+1)\theta$$

Comme $\sin(n+1)\theta_k = \sin(2k+1)\frac{\pi}{2} = (-1)^k$, il vient :

$$t'_{n+1}(u_k) = (n+1) \frac{(-1)^k}{\sin \theta_k}$$

D'où, en utilisant l'expression pour $u \neq u_k$, $l_k(u) = \frac{\pi_{n+1}(u)}{(u-u_k)\pi'_{n+1}(u_k)}$ et le fait que $\pi_{n+1} = 2^n t_{n+1}$:

$$\begin{aligned} \text{pour } \theta \neq \theta_k, \quad |l_k(\cos \theta)| &= \left| \frac{t_{n+1}(u)}{(u-u_k)t'_{n+1}(u_k)} \right| = \left| \frac{(-1)^k \sin \theta_k \cos(n+1)\theta}{(n+1)(\cos \theta - \cos \theta_k)} \right| \\ &= \frac{|\sin \theta_k \cos(n+1)\theta|}{(n+1)|\cos \theta - \cos \theta_k|} = \frac{|\sin \theta_k \cos(n+1)\theta|}{(n+1)|2 \sin(\frac{\theta-\theta_k}{2}) \sin(\frac{\theta+\theta_k}{2})|} \end{aligned}$$

Majoration : Pour $t \in [0, \frac{\pi}{2}]$, on a $\sin t \geq \frac{2}{\pi}t$.

Or, $\frac{\theta - \theta_k}{2} \in [-\frac{\pi}{2}, +\frac{\pi}{2}]$, donc $\left| \sin \frac{\theta - \theta_k}{2} \right| \geq \frac{|\theta - \theta_k|}{\pi}$.

Par ailleurs, $\frac{\theta + \theta_k}{2} \in [\frac{\theta_k}{2}, \frac{\theta_k + \pi}{2}]$, avec $\frac{\theta_k}{2} \leq \frac{\pi}{2}$ et $\frac{\theta_k + \pi}{2} \geq \frac{\pi}{2}$, donc

$$\left| \sin \frac{\theta + \theta_k}{2} \right| \geq \min\left\{ \sin \frac{\theta_k}{2}, \sin \frac{\theta_k + \pi}{2} \right\} = \min\left\{ \sin \frac{\theta_k}{2}, \cos \frac{\theta_k}{2} \right\}$$

Comme $\sin \theta_k = 2 \sin \frac{\theta_k}{2} \cos \frac{\theta_k}{2} \leq 2 \min\left\{ \sin \frac{\theta_k}{2}, \cos \frac{\theta_k}{2} \right\}$, on obtient, pour $\theta \neq \theta_k$:

$$|l_k(\cos \theta)| \leq \pi \frac{|\cos(n+1)\theta|}{(n+1)|\theta - \theta_k|}$$

On va utiliser cette majoration de deux façons différentes selon les valeurs de k : pour un θ fixé dans $[0, \pi]$, soit θ_j le point le plus proche de θ .

Pour $k = j, j+1, j-1$ on utilise le raisonnement suivant : par le théorème des accroissements finis, il existe $\xi \in [0, \pi]$ tel que

$$\cos(n+1)\theta = \cos(n+1)\theta - \cos(n+1)\theta_k = (n+1)(\theta - \theta_k)(-\sin \xi)$$

D'où,

$$|\cos(n+1)\theta| \leq (n+1)|\theta - \theta_k|$$

Donc $|l_k(\cos \theta)| \leq \pi$ pour tout $\theta \neq \theta_k$. Mais par continuité, ceci est encore vrai si $\theta = \theta_k$.

Pour $k \neq j, j+1, j-1$, on note $h = \frac{\pi}{n+1} = \theta_{k+1} - \theta_k$, alors on a :

$$|\theta - \theta_j| \leq \frac{h}{2}, \quad |\theta - \theta_k| \geq |\theta_j - \theta_k| - |\theta - \theta_j| \geq (|j - k| - 1)h$$

On en déduit : $|l_k(\cos \theta)| \leq \frac{\pi}{(n+1)h} \frac{1}{|j - k| - 1}$

D'où, en rassemblant les deux cas :

$$\sum_{k=0}^n |l_k(\cos \theta)| \leq \frac{\pi}{(n+1)h} \sum_{k \neq j, j+1, j-1} \frac{1}{|j - k| - 1} + 3\pi$$

d'où

$$\Lambda_n \leq 2\left(1 + \frac{1}{2} + \dots + \frac{1}{n}\right) + 3\pi \leq C \ln(n)$$

Minoration : Pour montrer l'inégalité inverse, on écrit :

$$\begin{aligned} \Lambda_n &\geq \sum_{k=0}^n |l_k(1)| = \frac{1}{n+1} \sum_{k=0}^n \left| \frac{(-1)^k \sin \theta_k}{(1 - \cos \theta_k)} \right| = \frac{1}{n+1} \sum_{k=0}^n \left| \frac{2 \sin \frac{\theta_k}{2} \cos \frac{\theta_k}{2}}{2 \sin \frac{\theta_k}{2} \sin \frac{\theta_k}{2}} \right| \\ &= \frac{1}{n+1} \sum_{k=0}^n \cotg \frac{\theta_k}{2} \geq \frac{2}{\pi} \int_{\frac{\theta_0}{2}}^{\frac{\pi}{2}} \cotgt \, dt = -\frac{2}{\pi} \ln\left(\sin \frac{\pi}{2(n+1)}\right) \geq \frac{2}{\pi} \ln(n) \end{aligned}$$

■

3

MÉTHODES D'APPROXIMATION
NUMÉRIQUE DES INTÉGRALES

Soit f une fonction définie sur un intervalle $[a, b]$. Pour définir l'intégrale de f sur $[a, b]$, on divise $[a, b]$ en n (n est quelconque), sous-intervalles $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ où $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$. Puis dans chaque sous-intervalle, on choisit un point $\zeta_1 \in [x_0, x_1], \zeta_2 \in [x_1, x_2], \dots, \zeta_n \in [x_{n-1}, x_n]$ et on calcule la somme

$$S(f) = f(\zeta_1)(x_1 - x_0) + f(\zeta_2)(x_2 - x_1) + \dots + f(\zeta_n)(x_n - x_{n-1})$$

Evidemment, cette somme dépend du choix de n , des points de subdivision x_1, x_2, \dots, x_{n-1} et des points $\zeta_1, \zeta_2, \dots, \zeta_n$.

Définition 3.0.1. On dit que f est intégrable sur $[a, b]$ si la somme $S(f)$ tend vers une limite quand n tend vers l'infini, quelque soit le choix des points de subdivision x_1, x_2, \dots, x_{n-1} et des points $\zeta_1, \zeta_2, \dots, \zeta_n$, pourvu que la longueur de chaque sous-intervalle tende vers zéro quand n tend vers l'infini. La limite de $S(f)$ s'appelle l'intégrale de Riemann de f sur $[a, b]$ et se note

$$\int_a^b f(x) dx.$$

On dit que x est la variable d'intégration et que $[a, b]$ est l'intervalle d'intégration.

On voit donc que $\int_a^b f(x) dx$ est la limite d'une somme et c'est cette notion qu'on doit garder à l'esprit quand on veut approcher une intégrale. Dans toute la suite, nous supposons que f est intégrable sur $[a, b]$. C'est le cas si, par exemple, f est continue sur $[a, b]$.

On connaît des formules qui permettent de calculer l'intégrale de quelques fonctions. Mais, contrairement à ce qui se passe pour le calcul des dérivées, il y a relativement peu de fonctions dont on sait calculer l'intégrale. Il suffit même de changer légèrement l'expression d'une fonction pour passer d'une fonction que l'on sait intégrer à une fonction qu'on ne sait pas intégrer. Par exemple, on sait que

$$\int_a^b \sin x dx = \cos a - \cos b,$$

mais on ne sait pas calculer

$$\int_a^b \frac{\sin x}{x} dx \quad \text{ou} \quad \int_a^b \sin(x^2) dx .$$

Or la fonction $\frac{\sin x}{x}$ est intégrable sur tout intervalle $[a, b]$ puisqu'elle est continue sur \mathbb{R} . Il en est de même pour la fonction $\sin(x^2)$. A défaut de pouvoir calculer ces intégrales, on doit savoir les approcher.

Il est donc important d'apprendre des techniques numériques pour approcher des intégrales. C'est indispensable lorsqu'il s'agit d'intégrales comme ci-dessus qu'on ne sait pas calculer explicitement.

Mais, même s'il s'agit d'intégrales qu'on sait calculer, leur calcul peut être long et compliqué et il peut être avantageux de le remplacer par un calcul approché.

Enfin, il y a un autre cas où il est indispensable d'approcher une intégrale : c'est lorsque l'expression de la fonction à intégrer n'est pas connue. Il se peut que la fonction ne soit connue qu'en un certain nombre de points : ses valeurs peuvent venir d'un autre calcul ou de mesures ponctuelles. Mais le cas le plus fréquent est celui où la fonction à intégrer fait partie des inconnues, par exemple lorsque c'est la solution d'une équation différentielle.

Un outil essentiel dans les méthodes d'approximation des intégrales est le résultat suivant :

Second Théorème de la Moyenne 3.0.2. *Soient f et g deux fonctions à valeurs réelles définies et continues sur un intervalle $[a, b]$. On suppose que g ne change pas de signe dans $[a, b]$. Alors, il existe un point c dans l'intervalle $[a, b]$ où on a l'égalité*

$$\int_a^b f(x) g(x) dx = f(c) \int_a^b g(x) dx .$$

Le premier théorème de la moyenne est un cas particulier de ce théorème ; il est obtenu en prenant $g = 1$.

Pour étudier une formule d'approximation d'une intégrale $\int_a^b f(x) dx$, on procédera en trois étapes :

- Définition d'une méthode élémentaire sur un intervalle de référence , par exemple $[0, 1]$ ou $[-1, +1]$.

-Passage à un intervalle $[x_0, x_0 + h]$ par changement de variable

-Découpage de l'intervalle $[a, b]$ en n sous intervalles $[x_i, x_i + h]$, $i = 0, 1, \dots, n$ sur lesquels on appliquera la méthode précédente pour obtenir une méthode composée sur $[a, b]$.

Les calculs d'erreur se feront également en suivant ces trois étapes.

3.1 Formules d'intégration à 1 point

Les formules d'intégration à un point sont les plus simples. On va en étudier deux, la formule des rectangles et la formule du point milieu.

Prenons par exemple l'intervalle $[0, 1]$ et soit f une fonction définie et continue sur $[0, 1]$.

Définition 3.1.1. On appelle formule des rectangles l'approximation suivante

$$\int_0^1 f(u) du \sim f(0).$$

Cette formule approche l'aire sous le graphe de f par l'aire du rectangle de base $[0, 1]$ et de hauteur $f(0)$. Calculons l'erreur que l'on commet quand on utilise cette formule, c'est-à-dire, calculons

$$\left| \int_0^1 f(u) du - f(0) \right|.$$

Nous avons vu au cours des chapitres précédents que les estimations d'erreurs demandent des hypothèses de régularité. Ce sera le cas aussi dans ce chapitre et nous ferons des hypothèses de régularité convenables au moment où nous en aurons besoin.

Proposition 3.1.2. Supposons que f soit dérivable et que sa dérivée soit continue sur $[0, 1]$. L'erreur dans la formule des rectangles est estimée par :

$$\int_0^1 f(u) du - f(0) = \frac{1}{2} f'(c)$$

où c est compris entre 0 et 1.

DÉMONSTRATION : On écrit

$$\int_0^1 f(u) du - f(0) = \int_0^1 (f(u) - f(0)) du$$

On peut aussi écrire

$$f(u) - f(0) = \int_0^u f'(t) dt.$$

Donc

$$\int_0^1 f(u) du - f(0) = \int_0^1 \left(\int_0^u f'(t) dt \right) du$$

Dans le système de coordonnées (u, t) , le domaine d'intégration est le triangle rectangle sous la droite $t = u$. Sans changer le domaine d'intégration, on peut changer l'ordre des variables d'intégration et on trouve

$$\int_0^1 \left(\int_0^u f'(t) dt \right) du = \int_0^1 \left(\int_t^1 f'(t) du \right) dt.$$

Mais

$$\int_t^1 f'(t) du = f'(t) \int_t^1 du = f'(t)(1 - t),$$

car $f'(t)$ ne dépend pas de u . Donc

$$\int_0^1 \left(\int_0^u f'(t) dt \right) du = \int_0^1 f'(t)(1-t) dt$$

Puisque la fonction $1-t$ est non négative dans l'intervalle $[0, 1]$, on peut appliquer le Second Théorème de la Moyenne, 3.0.2, avec $g(t) = 1-t$. Donc, il existe un point c dans l'intervalle $[0, 1]$ où on a l'égalité

$$\int_0^1 f'(t)(1-t) dt = f'(c) \int_0^1 (1-t) dt = \frac{1}{2} f'(c),$$

et on a trouvé une expression pour l'erreur :

$$\int_0^1 f(u) du - f(0) = \frac{1}{2} f'(c)$$

où c est un point dans l'intervalle $[0, 1]$, qui dépend de la fonction f . ■

A première vue ce résultat ne semble pas intéressant, mais c'est parce que l'intervalle $[0, 1]$ n'est pas petit. Supposons maintenant que l'intervalle d'intégration est $[x_0, x_0 + h]$, où h est un paramètre positif qu'on fera tendre vers zéro. Pour se ramener à une intégrale sur l'intervalle $[0, 1]$, on effectue le changement de variable :

$$x = x(u) = x_0 + h u.$$

Alors x décrit $[x_0, x_0 + h]$ quand u décrit $[0, 1]$ et

$$\int_{x_0}^{x_0+h} f(x) dx = \int_0^1 f(x_0 + h u) h du,$$

soit en notant $\hat{f}(u) = f(x_0 + h u)$,

$$\int_{x_0}^{x_0+h} f(x) dx = h \int_0^1 \hat{f}(u) du$$

En appliquant la formule des rectangles, et en remarquant que $\hat{f}(0) = f(x_0)$, on trouve la formule d'intégration

$$\int_{x_0}^{x_0+h} f(x) dx \sim h f(x_0)$$

Donc pour calculer l'erreur de cette formule, il ne reste qu'à calculer la dérivée de la fonction composée \hat{f} :

$$\hat{f}'(u) = h f'(x_0 + h u).$$

Avec ceci, l'estimation de 3.1.2 donne

$$\int_{x_0}^{x_0+h} f(x) dx - h f(x_0) = \frac{1}{2} h^2 f'(x_0 + h c)$$

où c est compris entre 0 et 1.

D'après cette formule, on voit que l'erreur tend vers zéro quand h tend vers zéro. Pour que cette estimation soit utilisable, il faut que la dérivée de f soit continue et bornée par une constante raisonnable.

Supposons maintenant que l'intervalle d'intégration d'origine est l'intervalle $[a, b]$. On suppose que $[a, b]$ est divisé en sous-intervalles égaux, c'est-à-dire

$$x_0 = a, \quad x_1 = a + h, \quad x_2 = a + 2h, \quad \dots, \quad x_{n-1} = a + (n-1)h, \quad x_n = b = a + nh,$$

avec $h = \frac{b-a}{n}$.

En appliquant la formule d'intégration ci-dessus à chacun de ces sous-intervalles et en sommant, on trouve

$$\int_a^b f(x) dx \sim h\{f(a) + f(a+h) + f(a+2h) + \dots + f(a+(n-1)h)\}.$$

En comparant avec 3.0.1, on constate que cette formule correspond à une somme de Rieman associée à un choix particulier de subdivision de l'intervalle $[a, b]$ et à un choix particulier de point dans chaque sous-intervalle. Donc, comme f est intégrable sur $[a, b]$ puisqu'elle est continue, la somme $h\{f(a) + f(a+h) + f(a+2h) + \dots + f(a+(n-1)h)\}$ converge vers $\int_a^b f(x) dx$ lorsque $h \rightarrow 0$.

La formule d'erreur 3.1.2 va nous donner sa vitesse de convergence. Pour cela, on suppose que f est dérivable dans $[a, b]$ et que f' est continue et bornée, c'est-à-dire, il existe une constante M_1 telle que

$$\forall x \in [a, b], \quad |f'(x)| \leq M_1.$$

Alors en appliquant la formule d'erreur à chaque sous-intervalle et en sommant, on trouve

$$\left| \int_a^b f(x) dx - h\{f(a) + f(a+h) + f(a+2h) + \dots + f(a+(n-1)h)\} \right| \leq \frac{1}{2} h^2 n M_1.$$

Mais

$$nh = b - a = \text{la longueur de l'intervalle d'intégration}.$$

Donc

$$\left| \int_a^b f(x) dx - h\{f(a) + f(a+h) + f(a+2h) + \dots + f(a+(n-1)h)\} \right| \leq \frac{1}{2} (b-a) h M_1.$$

Ceci montre que l'erreur tend vers zéro à la même vitesse que h . Si l'intervalle d'intégration n'est pas trop grand et si la constante M_1 n'est pas trop grande non plus, la méthode des rectangles est une bonne approximation de l'intégrale.

Remarque 3.1.3. La formule d'erreur est obtenue sous la seule hypothèse que la dérivée première de f est continue et bornée. Cette hypothèse est plus faible que celle qui est utilisée au Chapitre précédent pour estimer l'erreur d'approximation d'une dérivée. Ceci vient du fait que l'intégration est une opération régularisante : d'une certaine manière elle gomme les irrégularités de la fonction.

On obtient une formule d'intégration qui a des propriétés tout à fait semblables en prenant l'approximation

$$\int_0^1 f(u) du \sim f(1).$$

Mais, il est plus intéressant de choisir ce qu'on appelle la formule du point milieu. Comme pour la méthode des rectangles, on définira cette méthode sur l'intervalle de référence $[0, 1]$ avant de passer à un intervalle quelconque par un changement de variable.

Définition 3.1.4. La formule du point milieu est définie par :

$$\int_0^1 f(u) du \sim f\left(\frac{1}{2}\right).$$

C'est encore une formule de rectangles, mais c'est une formule centrée, contrairement aux formules précédentes. Nous allons voir que, de même que pour l'approximation des dérivées, une formule centrée est plus précise qu'une formule non centrée. Pour étudier convenablement l'erreur de cette formule, nous aurons besoin de la formule de Taylor suivante :

Théorème 3.1.5. (Formule de Taylor avec reste intégral) On suppose que la fonction f est n fois dérivable dans l'intervalle $[a, b]$ et que $f^{(n)}$ est continue. Alors

$$f(b) = f(a) + (b-a)f'(a) + \frac{1}{2}(b-a)^2 f''(a) + \dots + \frac{1}{(n-1)!} \int_a^b (b-t)^{n-1} f^{(n)}(t) dt.$$

Noter que la formule évidente

$$f(b) - f(a) = \int_a^b f'(t) dt,$$

est un cas particulier de la formule de Taylor avec reste intégral avec $n = 1$.

Proposition 3.1.6. Soit f une fonction deux fois continûment dérivable. Il existe un point c dans $[0, 1]$ tel que

$$\int_0^1 f(u) du - f\left(\frac{1}{2}\right) = \frac{1}{24} f''(c).$$

DÉMONSTRATION : Pour calculer

$$\int_0^1 f(u) du - f\left(\frac{1}{2}\right) = \int_0^1 [f(u) - f\left(\frac{1}{2}\right)] du,$$

nous appliquons la formule de Taylor avec $n = 2$, $a = \frac{1}{2}$ et $b = u$:

$$f(u) - f\left(\frac{1}{2}\right) = \left(u - \frac{1}{2}\right) f'\left(\frac{1}{2}\right) + \int_{\frac{1}{2}}^u (u-t) f''(t) dt.$$

On obtient

$$\int_0^1 f(u) du - f\left(\frac{1}{2}\right) = \int_0^1 \left(u - \frac{1}{2}\right) f'\left(\frac{1}{2}\right) du + \int_0^1 \int_{\frac{1}{2}}^u (u-t) f''(t) dt du.$$

D'une part, il est facile de vérifier que la première intégrale est nulle. D'autre part, on décompose la deuxième intégrale en deux :

$$\int_0^1 \int_{\frac{1}{2}}^u (u-t) f''(t) dt du = \int_0^{\frac{1}{2}} \int_u^{\frac{1}{2}} (t-u) f''(t) dt du + \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^u (u-t) f''(t) dt du,$$

et dans chacune on interchange l'ordre d'intégration. On trouve ainsi

$$\begin{aligned} \int_0^{\frac{1}{2}} \int_u^{\frac{1}{2}} (t-u) f''(t) dt du &= \int_0^{\frac{1}{2}} \int_0^t (t-u) f''(t) du dt \\ &= \int_0^{\frac{1}{2}} f''(t) \frac{1}{2} t^2 dt, \\ \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^u (u-t) f''(t) dt du &= \int_{\frac{1}{2}}^1 \int_t^1 (u-t) f''(t) du dt \\ &= \int_{\frac{1}{2}}^1 f''(t) \frac{1}{2} (1-t)^2 dt. \end{aligned}$$

Donc

$$\int_0^1 f(u) du - f\left(\frac{1}{2}\right) = \int_0^{\frac{1}{2}} f''(t) \frac{1}{2} t^2 dt + \int_{\frac{1}{2}}^1 f''(t) \frac{1}{2} (1-t)^2 dt.$$

Soit g la fonction définie sur $[0, 1]$ par

$$g(t) = \frac{1}{2} t^2, \quad \text{si } 0 \leq t \leq \frac{1}{2} \quad \text{et} \quad g(t) = \frac{1}{2} (1-t)^2, \quad \text{si } \frac{1}{2} \leq t \leq 1.$$

Ceci permet d'écrire plus simplement

$$\int_0^1 f(u) du - f\left(\frac{1}{2}\right) = \int_0^1 f''(t) g(t) dt.$$

Remarquer que g est continue et non négative sur l'intervalle $[0, 1]$ et

$$\int_0^1 g(t) dt = \frac{1}{24}.$$

On peut donc appliquer le Deuxième Théorème de la Moyenne au produit $f'' g$: il existe un point c dans $[0, 1]$ tel que

$$\int_0^1 f(u) du - f\left(\frac{1}{2}\right) = \frac{1}{24} f''(c).$$

■

Ensuite, pour déduire de 3.1.4 une formule d'intégration sur l'intervalle $[x_0, x_0 + h]$, on fait le changement de variable qu'on a déjà fait ci-dessus :

$$x = x_0 + h u \quad \text{et} \quad \hat{f}(u) = f(x_0 + h u).$$

D'où, la formule du point milieu :

$$\int_{x_0}^{x_0+h} f(x) dx \sim h f\left(x_0 + \frac{h}{2}\right).$$

Pour calculer l'erreur de la formule du point milieu, il ne reste qu'à calculer la dérivée seconde de la fonction composée \hat{f} :

$$\hat{f}''(u) = h^2 f''(x_0 + h u).$$

Alors la formule 3.1.6 donne

$$\int_{x_0}^{x_0+h} f(x) dx - h f\left(x_0 + \frac{h}{2}\right) = \frac{1}{24} h^3 f''(x_0 + h c)$$

où c est compris entre 0 et 1.

Enfin, il reste à appliquer cette formule à chaque sous-intervalle de l'intervalle d'origine $[a, b]$. Comme ci-dessus, on utilise des sous-intervalles égaux et on trouve la formule :

$$\int_a^b f(x) dx \sim h \left\{ f\left(a + \frac{h}{2}\right) + f\left(a + h + \frac{h}{2}\right) + \dots + f\left(a + (n-1)h + \frac{h}{2}\right) \right\}.$$

On peut remarquer qu'ici aussi, l'approximation de l'intégrale est faite à l'aide d'une somme de Rieman donc la différence tend vers 0 lorsque $h \rightarrow 0$.

Pour estimer l'erreur de cette formule, on suppose que f est deux fois dérivable dans $[a, b]$ et que f'' est continue et bornée, c'est-à-dire, il existe une constante M_2 telle que

$$\forall x \in [a, b], \quad |f''(x)| \leq M_2.$$

Alors, en appliquant la formule 3.1.6 dans chaque sous-intervalle et en sommant, on trouve

$$\left| \int_a^b f(x) dx - h \left\{ f\left(a + \frac{h}{2}\right) + f\left(a + h + \frac{h}{2}\right) + \dots + f\left(a + (n-1)h + \frac{h}{2}\right) \right\} \right| \leq \frac{1}{24} h^3 n M_2.$$

Puis, en tenant compte du fait que $nh = b - a$, ceci donne

$$\left| \int_a^b f(x) dx - h \left\{ f\left(a + \frac{h}{2}\right) + f\left(a + h + \frac{h}{2}\right) + \dots + f\left(a + (n-1)h + \frac{h}{2}\right) \right\} \right| \leq \frac{1}{24} (b-a) h^2 M_2.$$

Ceci montre que l'erreur de la formule du point milieu tend vers zéro à la même vitesse que h^2 . Si l'intervalle d'intégration n'est pas trop grand et si la constante M_2 n'est pas trop grande, l'erreur de la formule du point milieu est donc meilleure que l'erreur de la formule des rectangles.

Remarque 3.1.7. De même qu'au Chapitre 2, il faut faire attention aux hypothèses de régularité des énoncés du Second Théorème de la Moyenne et de la Formule de Taylor avec reste intégral. Leurs conclusions peuvent être fausses, lorsqu'on les applique à des fonctions qui n'ont pas la régularité requise par les hypothèses.

3.2 Une formule d'intégration à 2 points

La formule que nous proposons ici, la formule des trapèzes, est définie en intégrant, dans chaque sous-intervalle le polynôme d'interpolation de degré 1 prenant les mêmes valeurs que f aux extrémités de l'intervalle au lieu de la fonction. Comme dans le paragraphe précédent, on définit la formule sur l'intervalle $[0, 1]$, qui sert de référence, et on passe à un intervalle quelconque en faisant un changement de variable.

Soit donc f une fonction définie et continue sur $[0, 1]$ et soit p_1 son polynôme d'interpolation de degré un qui prend les mêmes valeur que f aux points 0 et 1 :

$$p_1(u) = f(0) + (f(1) - f(0))u.$$

Ainsi, comme

$$\int_0^1 p_1(u) du = \frac{1}{2}(f(0) + f(1)),$$

on propose la formule d'intégration

$$\int_0^1 f(u) du \sim \frac{1}{2}(f(0) + f(1)).$$

Définition 3.2.1. On appelle *Formule des Trapèzes*, la formule d'approximation :

$$\int_0^1 f(u) du \sim \frac{1}{2}(f(0) + f(1)).$$

Elle est appelée ainsi parce qu'elle approche l'aire sous la courbe par l'aire du trapèze de bases $[0, 1]$ et le segment $[(0, f(0)), (1, f(1))]$.

On peut calculer l'erreur de cette formule en appliquant la technique du paragraphe précédent. Mais les calcul sont plus compliqués et pour simplifier, on ne cherchera pas une égalité : on se contentera d'une borne supérieure. Pour cela, il est commode de se souvenir de la formule d'erreur du chapitre 2, Proposition 2.5.9, appliquée à l'ordre $n = 1$ et valable lorsque la fonction f est deux fois dérivable et f'' est continue :

Proposition 3.2.2. Si f'' est continue dans $[0, 1]$, alors

$$\left| \int_0^1 f(u) du - \frac{1}{2}(f(0) + f(1)) \right| \leq \frac{1}{12} \sup_{u \in [0, 1]} |f''(u)|.$$

DÉMONSTRATION : D'après la Proposition 2.5.9, si f'' est continue dans $[0, 1]$, alors pour chaque u dans $[0, 1]$, il existe un point c dans $[0, 1]$ où on a l'égalité

$$f(x) - p_1(x) = \frac{1}{2}x(x-1)f''(c).$$

En particulier, si f'' est bornée dans $[0, 1]$, on a

$$|f(u) - p_1(u)| \leq \frac{1}{2}u(1-u) \sup_{u \in [0, 1]} |f''(u)|,$$

car évidemment, $u(1-u) \geq 0$ dans $[0, 1]$.

Maintenant, calculons

$$\int_0^1 f(u) du - \frac{1}{2}(f(0) + f(1)).$$

Par construction

$$\int_0^1 f(u) du - \frac{1}{2}(f(0) + f(1)) = \int_0^1 (f(u) - p_1(u)) du.$$

Donc

$$\left| \int_0^1 f(u) du - \frac{1}{2}(f(0) + f(1)) \right| \leq \frac{1}{2} \sup_{u \in [0,1]} |f''(u)| \int_0^1 u(1-u) du,$$

c'est-à-dire

$$\left| \int_0^1 f(u) du - \frac{1}{2}(f(0) + f(1)) \right| \leq \frac{1}{12} \sup_{u \in [0,1]} |f''(u)|.$$

■

Ensuite, nous passons à un intervalle quelconque de la forme $[x_0, x_0 + h]$.

Rappelons le changement de variable

$$x = x_0 + hu \quad \text{et} \quad \hat{f}(u) = f(x_0 + hu).$$

Donc, sachant que

$$\int_{x_0}^{x_0+h} f(x) dx = h \int_0^1 \hat{f}(u) du,$$

et

$$\hat{f}(0) = f(x_0), \quad \hat{f}(1) = f(x_0 + h),$$

la Formule des Trapèzes sur $[x_0, x_0 + h]$ s'écrit

$$\int_{x_0}^{x_0+h} f(x) dx \sim \frac{1}{2}h(f(x_0) + f(x_0 + h)).$$

Pour calculer l'erreur de cette formule, rappelons que

$$\hat{f}''(u) = h^2 f''(x_0 + hu).$$

Alors 3.2.2 entraîne que

$$\left| \int_{x_0}^{x_0+h} f(x) dx - \frac{1}{2}h(f(x_0) + f(x_0 + h)) \right| \leq \frac{1}{12}h^3 \sup_{x \in [x_0, x_0+h]} |f''(x)|.$$

En comparant avec les résultats obtenus pour la formule du point milieu, nous voyons que la formule des trapèzes converge vers l'intégrale à la même vitesse que la formule du point milieu à un coefficient multiplicatif près.

Enfin, appliquons cette formule à chaque sous-intervalle de l'intervalle d'origine $[a, b]$. Comme au paragraphe précédent, on utilise n sous-intervalles égaux de longueur h , on applique la formule des trapèzes dans chaque sous-intervalle, on somme et on trouve

$$\int_a^b f(x) dx \sim h \left\{ \frac{1}{2} f(a) + f(a+h) + f(a+2h) + \dots + f(a+(n-1)h) + \frac{1}{2} f(b) \right\}.$$

Pour estimer l'erreur de cette formule, on suppose que f est deux fois dérivable dans $[a, b]$, que f'' est continue et qu'il existe une constante M_2 telle que

$$\forall x \in [a, b], |f''(x)| \leq M_2.$$

Alors, en appliquant la formule d'erreur 3.2.2 dans chaque sous-intervalle et en sommant, on trouve

$$\begin{aligned} \left| \int_a^b f(x) dx - h \left\{ \frac{1}{2} f(a) + f(a+h) + f(a+2h) + \dots + f(a+(n-1)h) + \frac{1}{2} f(b) \right\} \right| \\ \leq \frac{1}{12} h^3 n M_2 = \frac{1}{12} h^2 (b-a) M_2. \end{aligned}$$

3.3 Méthodes générales de quadrature

Soit $f : [a, b] \rightarrow \mathbb{R}$, une fonction continue. Dans cette partie, on va étudier des méthodes élémentaires de quadrature, à $k+1$ points, définies d'abord sur l'intervalle de référence $[-1, +1]$ et ensuite par changement de variable sur un intervalle $[x_0, x_0+h]$.

Ensuite, conformément au principe général, on va approcher l'intégrale $\int_a^b f(x) dx$ en effectuant un découpage de $[a, b]$ en n sous intervalles $[a_i, a_{i+1}]$, avec $a = a_0 < a_1 < \dots < a_n = b$, sur lesquels on appliquera des méthodes élémentaires.

Définition 3.3.1. On appelle méthode de quadrature élémentaire à $k+1$ points sur $[-1, +1]$, toute approximation de l'intégrale d'une fonction f sur $[-1, +1]$ par :

$$\int_{-1}^{+1} f(u) du \sim 2 \sum_{j=0}^k \omega_j f(c_j)$$

où $c_j \in [-1, +1]$, $0 \leq j \leq k$ et $\sum_{j=0}^k \omega_j = 1$. L'expression $2 \sum_{j=0}^k \omega_j f(c_j)$ est appelée

approximation quadratique élémentaire à $k+1$ points de l'intégrale $\int_{-1}^{+1} f(u) du$.

Proposition 3.3.2. La méthode de quadrature élémentaire à $k+1$ points sur $[x_0, x_0+h]$, associée à la méthode élémentaire de la Définition 3.3.1 est :

$$\int_{x_0}^{x_0+h} f(x) dx \sim h \sum_{j=0}^k \omega_j f(\xi_j)$$

où $\xi_j \in [x_0, x_0 + h]$, $0 \leq j \leq k$ et $\sum_{j=0}^k \omega_j = 1$.

DÉMONSTRATION : Il suffit d'effectuer le changement de variable $x = x_0 + \frac{h}{2} + \frac{h}{2}u$ et de poser $\xi_j = x_0 + \frac{h}{2} + \frac{h}{2}c_j$ pour $j = 0, \dots, k$. ■

Une façon de réaliser une méthode élémentaire est d'utiliser les polynômes d'interpolation d'ordre k de f aux points c_j , $j = 0, 1, \dots, k$, définis en 2.5.2 :

Proposition 3.3.3. Soient $c_j \in [-1, +1]$, $0 \leq j \leq k$, $k + 1$ points de l'intervalle $[-1, +1]$ et soit p_k le polynôme d'interpolation de f en ces points. Alors : $\int_{-1}^{+1} p_k(u) du$ définit une approximation quadratique élémentaire de f sur $[-1, +1]$.

DÉMONSTRATION : Par définition du polynôme d'interpolation de f , si l_j , $j = 0, 1, \dots, k$ sont les polynômes de base associés aux points c_j , $j = 0, 1, \dots, k$, on a :

$$\int_{-1}^{+1} p_k(u) du = \int_{-1}^{+1} \sum_{j=0}^k f(c_j) l_j(u) du = \sum_{j=0}^k f(c_j) \int_{-1}^{+1} l_j(u) du$$

L'intégrale de p_k sur $[-1, +1]$ est donc bien une expression de la forme recherchée pour une méthode de quadrature élémentaire, avec $\omega_j = \frac{1}{2} \int_{-1}^{+1} l_j(u) du$ si l'on vérifie que

$\sum_{j=0}^k \omega_j = 1$: on remarque que $\sum_{j=0}^k l_j(u) = 1$ car $\sum_{j=0}^k l_j(u)$ un polynôme de degré inférieur ou égal à k qui prend la valeur 1 aux $k + 1$ points distincts c_j , $j = 0, 1, \dots, k$. Par unicité du polynôme d'interpolation, ce polynôme coïncide donc avec le polynôme constant égal à 1.

Donc $\int_{-1}^{+1} \sum_{j=0}^k l_j(u) du = 2 = \sum_{j=0}^k \int_{-1}^{+1} l_j(u) du = 2 \sum_{j=0}^k \omega_j$, ce qui prouve le résultat. ■

Définition 3.3.4. On appelle méthode de quadrature composée toute approximation de l'intégrale d'une fonction f sur $[a, b]$ par :

$$\int_a^b f(x) dx \sim \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=0}^{k_i} \omega_{i,j} f(\xi_{i,j})$$

où $a = a_0 < a_1 < \dots < a_n = b$ et sur chaque intervalle $[a_i, a_{i+1}]$ l'expression $(a_{i+1} - a_i) \sum_{j=0}^{k_i} \omega_{i,j} f(\xi_{i,j})$ est une approximation quadratique élémentaire à $k_i + 1$ points de f

sur cet intervalle. L'expression $\sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=0}^{k_i} \omega_{i,j} f(\xi_{i,j})$ est appelée approximation

quadratique composée de l'intégrale $\int_a^b f(x) dx$.

A priori, on n'est pas obligé de prendre la même méthode de quadrature élémentaire sur chaque sous-intervalle $[a_i, a_{i+1}]$, bien que dans la pratique, cela soit souvent plus simple.

Définition 3.3.5. On dit qu'une méthode de quadrature (élémentaire ou composée) est d'ordre N si la formule approchée est exacte pour tout $p \in \mathcal{P}_N$ et inexacte pour au moins un $p \in \mathcal{P}_{N+1}$.

Remarquons que si $f \in \mathcal{P}_k$, la méthode de quadrature élémentaire, consistant à prendre l'intégrale du polynôme d'interpolation d'ordre k de f sur l'intervalle considéré, comme approximation quadratique de l'intégrale de f est d'ordre supérieur ou égale à k puisque dans le cas où f est un polynôme de degré k , cette fonction coïncide avec son polynôme d'interpolation d'ordre k .

Exemples 3.3.6. 1) Le cas où $k = 1$ et où l'approximation élémentaire est faite à l'aide du polynôme d'interpolation de f d'ordre 0 correspond à la méthode des rectangles ou à la méthode du point milieu selon le choix du point c , étudiées au paragraphe 1. On vérifie que la méthode des rectangles est d'ordre 0 et que la méthode du point milieu est d'ordre 1.

2) Le cas où $k = 2$ et où l'approximation élémentaire est faite à l'aide du polynôme d'interpolation de f d'ordre 1 avec des points équidistants correspond à la méthode des trapèzes étudiée au paragraphe 2. On vérifie que cette méthode est d'ordre 1.

3) Plus généralement, la méthode de Newton-Cotes d'ordre k consiste à faire l'approximation élémentaire à l'aide du polynôme d'interpolation de f d'ordre k , les points c_j , $j = 0, 1, \dots, k$ étant équidistants. On peut montrer que cette méthode est d'ordre k si k est impair et d'ordre $k + 1$ si k est pair. Sauf le cas $k = 1$, on utilisera ces méthodes uniquement dans le cas où k est pair.

Dans le cas 3), étudions les méthodes élémentaires de Newton-cotes sur l'intervalle de référence $[-1, +1]$:

* $k = 1$ on retrouve la méthode des trapèzes.

* $k = 2$ On a donc les points

$$c_0 = -1, \quad c_1 = 0, \quad c_2 = 1$$

et les polynômes de base

$$\forall u \in [-1, 1], \quad l_0(u) = \frac{u(u-1)}{2}, \quad l_1(u) = 1 - u^2, \quad l_2(u) = \frac{u(u+1)}{2}$$

D'où

$$\omega_0 = \omega_2 = \frac{1}{6}, \quad \omega_1 = \frac{2}{3}$$

Cette méthode s'appelle la méthode de Simpson et elle est d'ordre 3.

* $k = 4$ On trouve $\omega_0 = \omega_4 = \frac{7}{90}$, $\omega_1 = \omega_3 = \frac{16}{45}$, $\omega_2 = \frac{2}{15}$. La méthode est d'ordre 5 et s'appelle méthode de Boole-Villarceau.

* $k = 6$ On trouve $\omega_0 = \omega_6 = \frac{41}{840}$, $\omega_1 = \omega_5 = \frac{9}{35}$, $\omega_2 = \omega_4 = \frac{9}{280}$, $\omega_3 = \frac{34}{105}$. La méthode est d'ordre 7 et s'appelle méthode de Weddle-Hardy.

Pour $k \geq 8$, il apparaît des coefficients $\omega_j < 0$, ce qui rend les formules plus sensibles aux erreurs d'arrondis. Dans la pratique, on n'utilise ces méthodes que dans les quatre cas ci-dessus.

Dans certains cas simples, on peut démontrer que les approximations quadratiques convergent vers l'intégrale de la fonction :

Théorème 3.3.7. *Supposons que la méthode de quadrature élémentaire sur chaque intervalle $[a_i, a_{i+1}]$ comprenne un nombre $k_i = k$ fixe de points et que les coefficients $\omega_{i,j} = \omega_j$ ne dépendent pas de i . Alors, lorsque f est continue, l'approximation quadratique donnée par la méthode composée soit*

$$T_n(f) = \sum_{i=0}^{n-1} (a_{i+1} - a_i) \sum_{j=0}^k \omega_j f(\xi_{i,j})$$

converge vers $\int_a^b f(x) dx$ quand $n \rightarrow +\infty$

et quand le maximum du pas $h = \max\{(a_{i+1} - a_i) \mid i = 0, 1, \dots, n - 1\}$ tend vers 0.

DÉMONSTRATION : On peut écrire :

$$T_n(f) = \sum_{j=0}^k \omega_j S_{j,n}(f)$$

où $S_{j,n}(f) = \sum_{i=0}^{n-1} (a_{i+1} - a_i) f(\xi_{i,j})$ est une somme de Riemann relative à la subdivision $a_0 = a < a_1 < a_2 < \dots < a_n = b$.

Si f est continue, $S_{j,n}(f)$ converge vers $\int_a^b f(x) dx$ quand $n \rightarrow +\infty$ et quand le maximum du pas $h = \max\{(a_{i+1} - a_i) \mid i = 0, 1, \dots, n\}$ tend vers 0 et donc $T_n(f)$ aussi. ■

On cherche maintenant à évaluer l'erreur d'approximation dans une méthode de quadrature d'ordre N :

Définition 3.3.8. *On appelle erreur associée à la méthode de quadrature (élémentaire ou composée) $\int_a^b f(x) dx \sim \sum_{j=0}^k \lambda_j f(\xi_j)$ la différence*

$$E(f) = \int_a^b f(x) dx - \sum_{j=0}^k \lambda_j f(\xi_j)$$

Le résultat suivant est immédiat :

Propriété 3.3.9. *L'opérateur E défini ci-dessus, de $\mathcal{C}([a, b])$ dans \mathbb{R} est linéaire et vérifie :*

$$|E(f)| \leq [(b - a) + \sum_{j=0}^k |\lambda_j|] \sup_{x \in [a, b]} |f(x)|$$

On peut obtenir de meilleures estimations dans le cas d'une méthode de quadrature d'ordre N , pour des fonctions de classe C^{N+1} : pour cela, comme dans les paragraphes 1 et 2, pour une fonction de classe C^{N+1} sur $[a, b]$, on va utiliser la formule de Taylor avec reste intégral, 3.1.5 à l'ordre N :

$$f(x) = \sum_{k=0}^N \frac{1}{k!} f^{(k)}(a)(x-a)^k + \int_a^x \frac{1}{N!} (x-t)^N f^{(N+1)}(t) dt$$

et la deuxième formule de la moyenne, 3.0.2

$$\int_a^b f(x)w(x) dx = f(\xi) \int_a^b w(x) dx$$

où ξ est un point de $[a, b]$.

Par commodité, on va introduire une notation :

Notation. Soit x un réel, on pose

$$x_+ = \max(x, 0)$$

Il est clair que si $x > 0$, $x_+ = x$ et si $x \leq 0$, $x_+ = 0$.

La formule de Taylor avec reste intégral peut alors se ré-écrire :

$$f(x) = \sum_{k=0}^N \frac{1}{k!} f^{(k)}(a)(x-a)^k + \int_a^b \frac{1}{N!} (x-t)_+^N f^{(N+1)}(t) dt$$

Théorème 3.3.10. Supposons que la méthode de quadrature est d'ordre $N \geq 0$.

Si f est de classe C^{N+1} sur $[a, b]$, alors :

$$E(f) = \frac{1}{N!} \int_a^b K_N(t) f^{(N+1)}(t) dt$$

où K_N est une fonction sur $[a, b]$, appelée noyau de Peano associé à la méthode, défini par :

$$K_N(t) = E(x \rightarrow (x-t)_+^N), \quad t \in [a, b].$$

DÉMONSTRATION : Comme $f \rightarrow E(f)$ est une forme linéaire sur $\mathcal{E}[a, b]$, si g est une fonction intégrable de 2 variables $(x, t) \in [a, b] \times I$, le théorème de Fubini implique que

$$E\left(x \rightarrow \int_I g(x, t) dt\right) = \int_I E(x \rightarrow g(x, t)) dt$$

La formule de Taylor avec reste intégral rapellée ci-dessus donne l'existence d'un polynôme p de degré N tel que :

$$f(x) = p(x) + \int_a^b \frac{1}{N!} (x-t)_+^N f^{(N+1)}(t) dt$$

Comme $p \in \mathcal{P}_N$, on a $E(p) = 0$ par hypothèse, d'où

$$\begin{aligned} E(f) &= E\left(x \rightarrow \int_a^b \frac{1}{N!} (x-t)_+^N f^{(N+1)}(t) dt\right) \\ &= \int_a^b E\left(x \rightarrow \frac{1}{N!} (x-t)_+^N f^{(N+1)}(t)\right) dt \\ &= \int_a^b \frac{1}{N!} f^{(N+1)}(t) E\left(x \rightarrow (x-t)_+^N\right) dt \\ &= \frac{1}{N!} \int_a^b K_N(t) f^{(N+1)}(t) dt \end{aligned}$$

■

De ce résultat, on déduit immédiatement l'estimation suivante de l'erreur :

Corollaire 3.3.11. *Si la méthode de quadrature est d'ordre $N \geq 0$ et si f est de classe C^{N+1} sur $[a, b]$, on a la majoration :*

$$|E(f)| \leq \frac{1}{N!} \sup_{x \in [a, b]} \left| f^{(N+1)}(x) \right| \int_a^b |K_N(t)| dt.$$

Dans un cas particulier, on obtient une majoration de l'erreur plus facile à évaluer :

Corollaire 3.3.12. *Si la méthode de quadrature est d'ordre $N \geq 0$, si f est de classe C^{N+1} sur $[a, b]$ et si de plus K_N est de signe constant, alors pour tout $f \in \mathcal{C}^{N+1}([a, b])$, il existe $\xi \in [a, b]$ tel que*

$$E(f) = \frac{1}{N!} f^{(N+1)}(\xi) \int_a^b K_N(t) dt.$$

De plus, $\int_a^b K_N(t) dt = \frac{1}{N+1} E(x \rightarrow x^{N+1})$, donc

$$E(f) = \frac{1}{(N+1)!} f^{(N+1)}(\xi) E(x \rightarrow x^{N+1}).$$

DÉMONSTRATION : La première égalité est une conséquence du théorème 3.3.10 et de la formule de la moyenne 3.0.2 appliquée aux fonctions $f^{(N+1)}$ et K_N (ou $-K_N$ si $K_N \leq 0$). La deuxième égalité s'obtient en prenant $f(x) = x^{N+1}$ qui donne $f^{(N+1)}(x) = (N+1)!$. La troisième découle des deux premières. ■

Dans les exemples que nous avons étudiés, on peut évaluer l'erreur en utilisant le corollaire 3.3.12 :

Exemple 3.3.13. *Méthode du point milieu sur $[-1, +1]$.*

On a

$$E(f) = \int_{-1}^1 f(u) du - 2f(0)$$

Cette méthode est d'ordre 1 et le noyau de Peano est donné par :

$$\begin{aligned} K_1(t) &= E(u \rightarrow (u-t)_+) = \int_{-1}^1 (u-t)_+ du - 2(-t)_+ \\ &= \int_t^1 (u-t) du - 2(-t)_+ = \left[\frac{1}{2}(u-t)^2 \right]_t^1 - 2(-t)_+ = \frac{1}{2}(1-t)^2 - 2(-t)_+. \end{aligned}$$

On a donc

$$K_1(t) = \begin{cases} \frac{1}{2}(1-t)^2 & \text{si } t \geq 0 \\ \frac{1}{2}(1+t)^2 & \text{si } t \leq 0 \end{cases}$$

soit $K_1(t) = (\frac{1}{2}(1-|t|))^2$ sur $[-1, 1]$. Comme K_1 est positif et que

$$\int_{-1}^1 K_1(t) dt = \int_0^1 (1-t)^2 dt = \frac{1}{3}$$

le corollaire 3.3.12 implique

$$E(f) = \frac{1}{3} f''(\xi), \quad \xi \in]-1, 1[.$$

Exemple 3.3.14. *Méthode des trapèzes sur $[-1, +1]$.*

Cette méthode est aussi d'ordre 1 et on a

$$E(f) = \int_{-1}^1 f(u) du - [f(-1) + f(1)]$$

D'où

$$\begin{aligned} K_1(t) &= \int_{-1}^1 (u-t)_+ du - [(-1-t)_+ + (1-t)_+] \\ &= \int_t^1 (u-t) du - (1-t) \end{aligned}$$

$$K_1(t) = -\left[\frac{1}{2}(1-t^2) \right] \leq 0, \text{ sur } [-1, 1]$$

Comme $\int_{-1}^1 K_1(t) dt = -\frac{2}{3}$, on en déduit

$$E(f) = -\frac{2}{3} f''(\xi), \quad \xi \in]-1, 1[.$$

Exemple 3.3.15. Méthode de Simpson sur $[-1, +1]$.

C'est une méthode d'ordre 3 avec

$$E(f) = \int_{-1}^1 f(u) du - 2\left[\frac{1}{6}f(-1) + \frac{2}{3}f(0) + \frac{1}{6}f(1)\right]$$

D'où

$$\begin{aligned} K_3(t) &= E(x \rightarrow (u-t)_+^3) \\ &= \int_{-1}^1 (u-t)_+^3 du - 2\left[0 + \frac{2}{3}(-t)_+^3 + \frac{1}{6}(1-t)_+^3\right] \\ &= \int_t^1 (u-t)^3 du - 2\left[\frac{2}{3}(t)_-^3 + \frac{1}{6}(1-t)^3\right] \end{aligned}$$

D'où

$$\begin{aligned} K_3(t) &= \left[\frac{1}{4}(1-t)^4 - \frac{1}{3}(1-t)^3\right] = -\frac{1}{12}(1-t)^3(1+3t), \text{ sur } [0, 1] \\ K_3(t) &= -\frac{1}{12}(1-t)^3(1+3t) + \frac{4}{3}t^3 = -\frac{1}{12}(1+t)^3(1-3t), \text{ sur } [-1, 0] \end{aligned}$$

On a donc

$$K_3(t) = -\frac{1}{12}(1-|t|)^3(1+3|t|), \text{ sur } [-1, 1]$$

Comme K_3 garde un signe constant et que $\int_{-1}^1 K_3(t) dt = \frac{1}{4}E(x^4) = -\frac{1}{15}$, on en déduit

$$E(f) = -\frac{1}{15 \cdot 3!} f^{(4)}(\xi), \quad \xi \in]-1, 1[.$$

Voici un résultat général sur les méthodes de Newton-Cotes, utile dans les applications numériques car il permet d'utiliser la majoration de l'erreur donnée par le corollaire 3.3.12. Nous admettrons ce théorème :

Théorème 3.3.16. Dans les méthodes de Newton-Cotes, le noyau de Peano est de signe constant.

Pour les applications numériques, il peut être utile de connaître le résultat suivant :

Proposition 3.3.17. Le noyau de Peano K_N d'une méthode élémentaire d'ordre N ,

$$\int_{-1}^1 f(u) du \sim 2 \sum_{j=0}^k w_j f(\xi_j)$$

est une fonction paire dès que N est impair et que $\omega_{k-j} = \omega_j$, $\xi_{k-j} = -\xi_j$ (points et coefficients répartis symétriquement autour de 0).

DÉMONSTRATION : On commence par vérifier l'identité suivante, valable lorsque N est impair :

$$(u+t)_+^N - (-u-t)_+^N = (u+t)^N$$

puis on distingue k pair et k impair. ■

Nous allons maintenant étudier l'erreur de quadrature dans le cas d'une méthode composée associée à une méthode de quadrature élémentaire d'ordre N sur $[-1, +1]$:

$$\int_{-1}^1 g(u) du \sim 2 \sum_{j=0}^k w_j g(\tau_j)$$

Soit k_N le noyau de Peano associé à cette méthode élémentaire dont l'erreur correspondante est dénotée par :

$$E_e(g) = \int_{-1}^1 g(u) du - 2 \sum_{j=0}^k w_j g(\tau_j)$$

On considère un intervalle $[a, b]$ et une subdivision de cet intervalle :

$$a = a_0 < a_1 < \dots < a_n = b$$

Pour chaque $i = 0, 1, \dots, n-1$, posons $h_i = a_{i+1} - a_i$. On considère une méthode de quadrature composée associée à cette subdivision, telle que sur chaque sous intervalle $[a_i, a_{i+1}]$, la méthode élémentaire soit la méthode élémentaire définie ci-dessus, c'est à dire :

$$\int_a^b f(x) dx \sim \sum_{i=0}^{n-1} h_i \sum_{j=0}^k \omega_j f(\xi_{i,j})$$

où pour chaque $i = 0, 1, \dots, n-1$, les points $\xi_{i,j}$ se déduisent des τ_j par le changement de variable de $[-1, +1]$ dans $[a_i, a_{i+1}]$

$$u \rightarrow x = \frac{a_i + a_{i+1}}{2} + u \frac{h_i}{2}$$

On dénote l'erreur de la méthode composée par :

$$E_c(f) = \int_a^b f(x) dx - \sum_{i=0}^{n-1} h_i \sum_{j=0}^k \omega_j f(\xi_{i,j})$$

Pour chaque $i = 0, 1, \dots, n-1$, définissons $g_i \in \mathcal{C}([-1, +1])$ par :

$$g_i(u) = f\left(\frac{a_i + a_{i+1}}{2} + u \frac{h_i}{2}\right)$$

Comme sur chaque intervalle $[a_i, a_{i+1}]$, $dx = \frac{h_i}{2} du$, on peut calculer $E_c(f)$ en décomposant l'intégrale sur $[a, b]$ et en utilisant les g_i sur chaque $[a_i, a_{i+1}]$:

$$E_c(f) = \sum_{i=0}^{n-1} \left(\frac{h_i}{2} \int_{-1}^{+1} g_i(u) du - h_i \sum_{j=0}^k \omega_j g_i(\tau_j) \right) = \sum_{i=0}^{n-1} \frac{h_i}{2} E_e(g_i)$$

Le noyau de Peano de la méthode composée est donc, pour $t \in [a, b]$:

$$\begin{aligned}
 K_N(t) &= E_c \left(x \rightarrow (x-t)_+^N \right) = \sum_{i=0}^{n-1} \frac{h_i}{2} E_e \left(u \rightarrow \left(\frac{a_i + a_{i+1}}{2} + u \frac{h_i}{2} - t \right)_+^N \right) \\
 &= \sum_{i=0}^{n-1} \frac{h_i}{2} E_e \left(u \rightarrow \left(\frac{h_i}{2} \right)^N \left(u - \frac{2}{h_i} \left(t - \frac{a_i + a_{i+1}}{2} \right) \right)_+^N \right) \\
 &= \sum_{i=0}^{n-1} \left(\frac{h_i}{2} \right)^{N+1} E_e \left(u \rightarrow \left(u - \frac{2}{h_i} \left(t - \frac{a_i + a_{i+1}}{2} \right) \right)_+^N \right)
 \end{aligned}$$

Pour chaque $i = 0, 1, \dots, n-1$, posons : $\theta_i(t) = \frac{2}{h_i} \left(t - \frac{a_i + a_{i+1}}{2} \right)$.

Supposons $t \in [a_j, a_{j+1}]$, alors :

1) $\theta_i(t) \in [-1, +1] \iff i = j$

2) Si $i \neq j$

i) ou bien $\theta_i(t) > 1$ et pour $u \in [-1, +1]$, $(u - \theta_i(t))_+^N = 0$

ii) ou bien $\theta_i(t) \leq -1$ et pour $u \in [-1, +1]$, $(u - \theta_i(t))_+^N = (u - \theta_i(t))^N$

Dans les deux cas 2)i) et 2)ii), $(u - \theta_i(t))_+^N$ est un polynôme de degré $\leq N$ sur $[-1, +1]$ donc $E_e \left(u \rightarrow (u - \theta_i(t))_+^N \right) = 0$.

Donc seul le terme en $i = j$ est non nul dans la sommation définissant K_N , d'où, pour $t \in [a_j, a_{j+1}]$:

$$\begin{aligned}
 K_N(t) &= \left(\frac{h_j}{2} \right)^{N+1} E_e \left(u \rightarrow \left(u - \frac{2}{h_j} \left(t - \frac{a_j + a_{j+1}}{2} \right) \right)_+^N \right) \\
 &= \left(\frac{h_j}{2} \right)^{N+1} k_N \left(\frac{2}{h_j} \left(t - \frac{a_j + a_{j+1}}{2} \right) \right)
 \end{aligned}$$

Théorème 3.3.18. Avec les notations définies ci-dessus, on suppose que k_N est de signe constant et que h_i est constant, égal à $h = \frac{b-a}{n}$. On note $C_N = \int_{-1}^{+1} k_N(t) dt$. Alors, pour tout $f \in \mathcal{C}^{N+1}([a, b])$, il existe un point $\xi \in]a, b[$ tel que

$$E_c(f) = \frac{C_N}{N! 2^{N+2}} h^{N+1} f^{(N+1)}(\xi) (b-a)$$

DÉMONSTRATION : K_N étant aussi de signe constant, le corollaire 3.3.12 montre l'existence d'un $\xi \in]a, b[$ tel que

$$E_c(f) = \frac{1}{N!} f^{(N+1)}(\xi) \int_a^b K_N(t) dt$$

D'après l'expression de K_N , on obtient :

$$\int_a^b K_N(t) dt = n \int_{a_0}^{a_1} K_N(t) dt = n \left(\frac{h}{2} \right)^{N+1} \int_{a_0}^{a_1} k_N \left(\frac{2}{h} \left(t - \frac{a_0 + a_1}{2} \right) \right) dt$$

En effectuant le changement de variable $t = \frac{a_0 + a_1}{2} + \frac{h}{2}u$, on obtient :

$$\int_a^b K_N(t) dt = n \left(\frac{h}{2}\right)^{N+2} \int_{-1}^{+1} k_N(u) du = nh \frac{h^{N+1}}{2^{N+2}} C_N = \frac{C_N}{2^{N+2}} h^{N+1} (b-a)$$

D'où, en remplaçant :

$$E_c(f) = \frac{C_N}{N! 2^{N+2}} h^{N+1} f^{(N+1)}(\xi)(b-a)$$

■

On voit donc que, si f est suffisamment régulière, lorsque le pas h tend vers 0, l'ordre de grandeur de l'erreur d'une méthode composée est h^{N+1} : il est donc intéressant d'utiliser des méthodes d'ordre élevé qui donnent une précision meilleure.

Exemples 3.3.19.

Point milieu : $N = 1$, $C_1 = \frac{1}{3}$, $E_c(f) = \frac{1}{24} h^2 f''(\xi)(b-a)$

Trapèzes : $N = 1$, $C_1 = -\frac{2}{3}$, $E_c(f) = -\frac{1}{12} h^2 f''(\xi)(b-a)$

Simpson : $N = 3$, $C_3 = -\frac{1}{15}$, $E_c(f) = -\frac{1}{2880} h^4 f^{(4)}(\xi)(b-a)$.

4

MÉTHODES D'APPROXIMATION NUMÉRIQUE DES SOLUTIONS D'ÉQUATIONS

On va présenter dans ce chapitre des méthodes itératives, permettant la résolution approchée d'une équation $f(x) = 0$, où f est une fonction d'un sous-ensemble de \mathbb{R}^m dans \mathbb{R}^m , $m \geq 1$.

Le principe consiste à remplacer cette équation par une équation équivalente du type $\varphi(x) = x$, plus facile à résoudre grâce à l'outil essentiel qu'est le Théorème du point fixe.

4.1 Principe des méthodes itératives

L'outil de base est un résultat classique de topologie, le théorème du point fixe :

Théorème 4.1.1. *Soit E un espace métrique complet et φ une application strictement contractante de E dans lui-même, c'est à dire qu'il existe $k \in]0, 1[$ tel que*

$$\forall x_1, x_2 \in E, \quad d(\varphi(x_1), \varphi(x_2)) \leq kd(x_1, x_2)$$

Alors φ admet un point fixe unique, c'est à dire qu'il existe $a \in E$ tel que $\varphi(a) = a$.

De plus, toute suite $(x_n)_{n \in \mathbb{N}}$ telle que $x_{n+1} = \varphi(x_n)$ converge vers le point fixe.

Exemple. *Tout espace vectoriel normé complet, toute partie fermée d'un espace vectoriel normé complet est un espace métrique complet. En particulier, \mathbb{R}^m est complet pour tout $m \in \mathbb{N}$.*

DÉMONSTRATION : Commençons par l'unicité : sous les hypothèses du théorème, s'il y avait deux points fixes $a \neq b$, alors $d(\varphi(a), \varphi(b)) = d(a, b)$. Comme $d(a, b) \neq 0$, φ ne pourrait pas être contractante. Il n'y a donc bien qu'un seul point fixe possible.

Soit x_0 un point quelconque de E et $(x_n)_{n \in \mathbb{N}}$ la suite récurrente définie dans le théorème, de point initial x_0 . Alors

$$d(x_p, x_{p+1}) = d(\varphi(x_{p-1}), \varphi(x_p)) \leq kd(x_{p-1}, x_p)$$

d'où par récurrence

$$d(x_p, x_{p+1}) \leq k^p d(x_0, x_1)$$

Pour tout entier $q > p$, il vient :

$$d(x_p, x_q) \leq \sum_{l=p}^{q-1} d(x_l, x_{l+1}) \leq \left(\sum_{l=p}^{q-1} k^l \right) d(x_0, x_1)$$

avec

$$\sum_{l=p}^{q-1} k^l \leq \sum_{l=p}^{\infty} k^l = \frac{k^p}{1-k}$$

On a donc,

$$d(x_p, x_q) \leq \frac{k^p}{1-k}$$

pour tout $p < q$.

La suite $(x_n)_{n \in \mathbb{N}}$ est donc une suite de Cauchy. Comme l'espace E est complet, cette suite converge vers un point $a \in E$. L'égalité $x_{n+1} = \varphi(x_n)$ et la continuité de φ impliquent à la limite que $a = \varphi(a)$. ■

Pour les approximations numériques, la vitesse de convergence de la suite $(x_n)_{n \in \mathbb{N}}$ vers le point fixe a joue un rôle essentiel. Elle est estimée par :

Proposition 4.1.2. *Sous les hypothèses du théorème 4.1.1, on a :*

$$d(x_p, a) \leq k^p d(x_0, a).$$

DÉMONSTRATION : Il suffit d'écrire $d(x_p, a) = d(\varphi(x_{p-1}), \varphi(a)) \leq kd(x_{p-1}, a)$ et par récurrence, on obtient l'inégalité cherchée. ■

La convergence est donc exponentiellement rapide, ce qui est très satisfaisant.

4.2 Utilisation du théorème du point fixe pour des fonctions réelles de classe C^1

Soit I un intervalle fermé de \mathbb{R} et $\varphi : I \rightarrow I$ une application de classe C^1 . Soit $a \in I$ un point fixe de φ . Distinguons trois cas :

Cas 1 : $|\varphi'(a)| < 1$.

Par continuité de φ' , il existe un voisinage $E = [a - h, a + h]$ de a sur lequel $|\varphi'(x)| \leq k < 1$. Remarquons d'abord que E est stable par φ , c'est à dire $\varphi(E) \subset E$, puisque $\varphi(a) = a$ et φ est contractante sur E . Le théorème des accroissements finis, 2.1.7, implique alors que sur E , on est dans le cadre d'application du théorème du point fixe, 4.1.1. On a donc convergence de la suite $(x_n)_{n \in \mathbb{N}}$ vers a avec vitesse exponentielle. Dans ce cas, le point fixe a est dit attractif.

Cas Particulier 4.2.1. *Dans le cas où $\varphi'(a) = 0$, φ de classe C^2 et $|\varphi''| \leq M$, la vitesse de convergence est de l'ordre de $|x_0 - a|^{2^p}$. On dit dans ce cas que l'on a une convergence quadratique et le point fixe est appelé super attractif.*

DÉMONSTRATION : On utilise la formule de Taylor à l'ordre 2 : il existe $c \in]a, x[$ tel que

$$\varphi(x) = \varphi(a) + (x - a)\varphi'(a) + \frac{(x - a)^2}{2}\varphi''(c) = a + \frac{(x - a)^2}{2}\varphi''(c)$$

D'où $|\varphi(x) - a| \leq \frac{1}{2}M|x - a|^2$, soit encore : $\frac{1}{2}M|\varphi(x) - a| \leq [M\frac{1}{2}|x - a|]^2$.

Par récurrence, on en déduit successivement

$$\frac{1}{2}M|x_p - a| \leq [M\frac{1}{2}|x_0 - a|]^{2^p}$$

$$|x_p - a| \leq \frac{2}{M}[M\frac{1}{2}|x_0 - a|]^{2^p}$$

■

Cas 2 : $|\varphi'(a)| > 1$

Comme

$$\lim_{x \rightarrow a} \left| \frac{\varphi(x) - \varphi(a)}{x - a} \right| = |\varphi'(a)| > 1$$

il existe un voisinage $E = [a - h, a + h]$ de a sur lequel

$$x \neq a, |\varphi(x) - a| > |x - a|$$

On dit alors que le point a est répulsif et la suite récurrente du théorème 4.1.1 ne converge pas

Par contre, quitte à diminuer h , d'après le théorème d'inversion locale, la restriction de φ à E est inversible et son inverse est de classe C^1 . Comme $(\varphi^{-1})'(a) = \frac{1}{\varphi'(a)}$, le point a est attractif pour φ^{-1} .

Cas 3 : $|\varphi'(a)| = 1$

On est ici dans un cas douteux, comme le montrent les deux exemples suivants pour lesquels $a = 0$ et $\varphi'(a) = 1$:

Exemple 4.2.2. On considère la fonction $\varphi(x) = \sin x$ sur $A = [0, \pi/2]$. Comme $0 < \sin x < x$ sur $]0, \pi/2]$, pour tout $x_0 \in E$, la suite des itérés est décroissante et minorée donc converge vers l . Comme $l = \sin l$, $l = 0$.

Exemple 4.2.3. On considère la fonction $\varphi(x) = \operatorname{sh} x$ sur $A = [0, +\infty[$. Comme $\operatorname{sh} x > x$ sur $]0, +\infty]$, pour tout $x_0 \in E$, on voit que $\lim_{p \rightarrow \infty} x_p = +\infty$ et que le point fixe $a = 0$ est répulsif.

On va maintenant traiter un exemple où on cherche à approcher la solution d'une équation réelle $f(x) = 0$ en remplaçant cette équation par une équation du type $\varphi(x) = x$, qui permet d'utiliser les résultats précédents :

Exemple 4.2.4. Résoudre l'équation $f(x) = x^3 - 4x + 1 = 0$ sur \mathbb{R} .

On a $f'(x) = 3x^2 - 4$, les zéros de la dérivée sont donc $\frac{-2}{\sqrt{3}}, \frac{2}{\sqrt{3}}$. En ces points, la fonction f prend les valeurs $1 + \frac{16}{3\sqrt{3}} > 0$ et $1 - \frac{16}{3\sqrt{3}} < 0$. Le tableau de variation de f et l'étude de quelques valeurs de f donnent facilement l'existence de 3 racines $\alpha_1 < \alpha_2 < \alpha_3$ de l'équation $f(x) = 0$ telles que

$$-2,5 < \alpha_1 < -2, \quad 0 < \alpha_2 < 0,5, \quad 1,5 < \alpha_3 < 2$$

L'équation $f(x) = 0$ peut se ré-écrire $\varphi(x) = x$, avec $\varphi(x) = \frac{1}{4}(x^3 + 1)$. On a alors $\varphi'(x) = \frac{3}{4}x^2$, d'où :

Pour $x \in [-2,5; -2]$, $\varphi'(x) \geq \varphi'(2) = 3$

Pour $x \in [0; 0,5]$, $0 \leq \varphi'(x) \leq 0,1875$

Pour $x \in [1,5; 2]$, $\varphi'(x) \geq \varphi'(1,5) = 1,6875$

Seul α_2 est un point fixe attractif de φ . L'intervalle $[0; 0,5]$ est nécessairement stable par φ puisqu'il contient un point fixe et que φ est contractante et croissante. On peut donc appliquer le théorème du point fixe sur cet intervalle et pour tout $x_0 \in [0; 0,5]$, on aura $\alpha_2 = \lim_{p \rightarrow +\infty} x_p$.

Pour obtenir α_1 et α_3 , on peut utiliser la fonction $\varphi^{-1}(x) = \sqrt[3]{4x - 1}$, qui est contractante au voisinage de ces points.

On peut remarquer qu'il est plus efficace du point de vue numérique d'écrire l'équation sous la forme $x^2 - 4 + \frac{1}{x} = 0$, soit $x = \varphi_+(x)$ ou $x = \varphi_-(x)$ avec :

$$\varphi_+(x) = \sqrt{4 - \frac{1}{x}}, \quad \varphi_-(x) = -\sqrt{4 - \frac{1}{x}}$$

pour $x \geq 0$ et $x \leq 0$ respectivement.

On a alors $\varphi_+(x) = \frac{1}{2x^2}(4 - \frac{1}{x})^{-1/2}$ et $\varphi_-(x) = -\frac{1}{2x^2}(4 - \frac{1}{x})^{-1/2}$, ce qui donne :

pour $x \in [1,5; 2]$, $0 \leq \varphi'_+(x) \leq \varphi'_+(1,5) \sim 0,122$

pour $x \in [-2,5; -2]$, $\varphi'_-(-2) \sim -0,059 \leq \varphi'_-(x) \leq 0$

et la convergence sera assez rapide. ■

La méthode de Newton, que nous allons définir dans le paragraphe suivant, donne un moyen systématique d'associer à une équation réelle $f(x) = 0$ donnée, une équation équivalente du type $\varphi(x) = x$, où φ est une application contractante d'une partie fermée $A \subset \mathbb{R}$ dans elle-même. Grâce au Théorème du point fixe, ceci permettra ensuite d'approcher la solution a de l'équation $f(x) = 0$ par une suite récurrente associée à φ .

4.3 Méthode de Newton pour les fonctions réelles

Soit f une fonction définie sur un intervalle de \mathbb{R} , à valeurs dans \mathbb{R} . On suppose qu'on possède une valeur grossière x_0 de la solution a de l'équation $f(x) = 0$. L'idée est de remplacer f par sa tangente au point x_0 , d'équation

$$y = f'(x_0)(x - x_0) + f(x_0)$$

Théorème 4.3.1. (Méthode de Newton) Si f est de classe C^2 et si $f'(a) \neq 0$, la racine a de l'équation $f(x) = 0$ est un point fixe superattractif pour la fonction φ , définie sur un intervalle $I = [a - r, a + r]$ et telle que

$$\varphi(x) = x - \frac{f(x)}{f'(x)}.$$

DÉMONSTRATION : L'intervalle $I = [a - r, a + r]$ est déterminé par la propriété :

$$\forall x \in I, f'(x) \neq 0$$

qui est assurée par l'hypothèse $f'(a) \neq 0$ et la continuité de f' .

L'abscisse x_1 du point d'intersection de la tangente à la courbe $y = f(x)$ au point x_0 avec l'axe des x est donnée par

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

x_1 est en général une meilleure approximation de a que x_0 . On va donc itérer la fonction

$$\varphi(x) = x - \frac{f(x)}{f'(x)}$$

Si f est de classe C^2 et si $f'(a) \neq 0$, φ est de classe C^1 et :

$$\varphi'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}$$

ce qui donne $\varphi(a) = a$ et $\varphi'(a) = 0$. La racine a de l'équation $f(x) = 0$ est donc bien un point superattractif pour φ . ■

Si la fonction φ est de classe C^2 , c'est à dire si f est de classe C^3 , le cas particulier 4.2.1 donne une évaluation de la vitesse de convergence vers le point fixe. Dans le cas plus général où f est seulement de classe C^2 , on a également cette estimation grâce au résultat suivant :

Théorème 4.3.2. Supposons que f est de classe C^2 sur $I = [a - r, a + r]$ et que $f' \neq 0$ sur I . Soit $M = \max_{x \in I} \left| \frac{f''(x)}{f'(x)} \right|$ et $h = \min(r, \frac{1}{M})$. Alors, pour tout $x \in [a - h, a + h]$, on a $|\varphi(x) - a| \leq M |x - a|^2$ et pour tout point initial $x_0 \in [a - h, a + h]$,

$$|x_p - a| \leq \frac{1}{M} (M |x_0 - a|)^{2^p}.$$

DÉMONSTRATION : On introduit la fonction $u(x) = \frac{f(x)}{f'(x)}$ qui vérifie donc $\varphi'(x) = u(x) \frac{f''(x)}{f'(x)}$. La fonction f étant monotone sur I et nulle en a par hypothèse, elle a le même signe que $f'(a)(x - a)$ ce qui entraîne que $u(x)$ a même signe que $x - a$. De plus :

$$u'(x) = 1 - \frac{f(x)f''(x)}{f'(x)^2} = 1 - \frac{f''(x)}{f'(x)}u(x)$$

donc $|u'(x)| \leq 1 + M |u(x)|$ sur I . On a besoin du lemme suivant :

Lemme. On a, pour $x \in I$,

$$|u(x)| \leq \frac{1}{M} (e^{M|x-a|} - 1).$$

Montrons cette inégalité par exemple pour $x \geq a$. On pose $v(x) = u(x)e^{-Mx}$, alors, comme dans ce cas, $u(x) \geq 0$, $u'(x) \leq 1 + Mu(x)$ et donc :

$$v'(x) = (u'(x) - Mu(x))e^{-Mx} \leq e^{-Mx}$$

Comme $v(a) = u(a) = 0$, on en déduit par intégration

$$v(x) \leq \frac{1}{M} (e^{-Ma} - e^{-Mx})$$

soit encore

$$u(x) \leq \frac{1}{M} (e^{M(x-a)} - 1).$$

■

On a besoin d'un deuxième lemme :

Lemme. Pour $|t| \leq 1$, $e^{|t|} - 1 \leq 2|t|$.

En effet, la fonction exponentielle est convexe, donc sur tout intervalle, la courbe est située sous sa corde. Sur l'intervalle $[0, 1]$, cela donne $e^t \leq 1 + (e - 1)t$, d'où le lemme puisque $e - 1 < 2$. ■

On revient à la démonstration du théorème : On écrit $\varphi'(x) = u(x) \frac{f''(x)}{f'(x)}$ ce qui donne grâce au premier lemme : $|\varphi'(x)| \leq M |u(x)| \leq e^{M|x-a|} - 1$. Grâce au deuxième lemme, on obtient $|\varphi'(x)| \leq 2M |x - a|$ pour $|x - a| \leq \min(r, \frac{1}{M})$. Comme $\varphi(a) = a$, on voit par intégration que pour $x \in [a - h, a + h]$, on a

$$|\varphi(x) - a| \leq 2M |x - a|^2$$

soit

$$2M |\varphi(x) - a| \leq (2M |x - a|)^2$$

On obtient ensuite l'inégalité cherchée par itération. ■

On a déjà vu au chapitre 2 que dans bien des cas, la dérivée d'une fonction ne peut être calculée explicitement. Dans un cas de ce genre, la méthode de Newton ne peut pas s'appliquer. C'est pourquoi nous introduisons une autre méthode, la méthode de la sécante, qui consiste à remplacer la dérivée par le taux d'accroissement de la fonction. La méthode ainsi obtenue ne relève pas directement du Théorème du point fixe et nécessite une démonstration un peu différente.

4.4 Méthode de la sécante pour les fonctions réelles

Soit f une fonction définie sur un intervalle de \mathbb{R} , à valeurs dans \mathbb{R} telle que $f(a) = 0$. On suppose que l'on a deux valeurs approchées x_0 et x_1 de la racine a , l'équation de la sécante passant par les points $(x_0, f(x_0))$ et $(x_1, f(x_1))$ est alors :

$$y = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_1) + f(x_1)$$

Définition 4.4.1. On pose

$$\tau_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

et on définit par récurrence

$$\tau_p = \frac{f(x_p) - f(x_{p-1})}{x_p - x_{p-1}}, \quad x_{p+1} = x_p - \frac{f(x_p)}{\tau_p}.$$

La convergence de la suite $(x_p)_{p \in \mathbb{N}}$ et sa vitesse de convergence sont données par le résultat suivant :

Théorème 4.4.2. (Méthode de la sécante) On suppose que f est de classe C^2 de dérivée $f' \neq 0$ sur $I = [a - r, a + r]$. Pour $i = 1, 2$, on pose $M_i = \max_{x \in I} |f^{(i)}(x)|$, $m_i = \min_{x \in I} |f^{(i)}(x)|$ et on introduit les constantes $K = \frac{M_2}{2m_1}(1 + \frac{M_1}{m_1})$ et $h = \min(r, \frac{1}{K})$. Alors, quelque soit le choix des points initiaux x_0 et x_1 dans $[a - h, a + h]$ distincts, on a :

$$|x_p - a| \leq \frac{1}{K} [K \max(|x_0 - a|, |x_1 - a|)]^{s_p}$$

où $(s_p)_{p \in \mathbb{N}}$ est la suite de Fibonacci, définie par $s_{p+1} = s_p + s_{p-1}$ et $s_0 = s_1 = 1$.

DÉMONSTRATION : On considère le taux d'accroissement $\tau(x, y)$ de f sur $I \times I$:

$$\begin{cases} \tau(x, y) = \frac{f(y) - f(x)}{y - x} & \text{si } y \neq x \\ \tau(x, x) = f'(x) \end{cases}$$

Pour tout $(x, y) \in I \times I$, on peut écrire :

$$\tau(x, y) = \int_0^1 f'(x + t(y - x)) dt$$

et le théorème de dérivation sous le signe somme implique :

$$\frac{\partial \tau}{\partial x}(x, y) = \int_0^1 (1 - t) f''(x + t(y - x)) dt$$

$$\frac{\partial \tau}{\partial y}(x, y) = \int_0^1 t f''(x + t(y - x)) dt$$

Comme f' est de signe constant sur I et comme $\int_0^1 (1 - t) dt = \int_0^1 t dt = \frac{1}{2}$, on en déduit :

$$|\tau(x, y)| \geq m_1, \quad \left| \frac{\partial \tau}{\partial x}(x, y) \right| \leq \frac{1}{2} M_2, \quad \left| \frac{\partial \tau}{\partial y}(x, y) \right| \leq \frac{1}{2} M_2$$

Il résulte en particulier que

$$|\tau(x, y) - f'(x)| = |\tau(x, y) - \tau(x, x)| = \left| \int_x^y \frac{\partial \tau}{\partial y}(x, t) dt \right| \leq \frac{1}{2} M_2 |y - x|$$

La suite $(x_p)_{p \in \mathbb{N}}$ est définie par la formule de récurrence $x_{p+1} = \psi(x_p, x_{p-1})$ où ψ est la fonction définie par

$$\psi(x, y) = x - \frac{f(x)}{\tau(x, y)}$$

En posant $h_p = x_p - a$, on a :

$$h_{p+1} = \psi(x_p, x_{p-1}) - a = \psi(a + h_p, a + h_{p-1}) - \psi(a, a)$$

En particulier $h_2 = \psi(x_1, x_0) - a = \psi(a + h_1, a + h_0) - \psi(a, a)$.

En intégrant sur $[0, 1]$ la dérivée de la fonction $t \rightarrow \psi(a + th_1, a + th_0)$, on trouve :

$$h_2 = \int_0^1 \left[h_1 \frac{\partial \psi}{\partial x}(a + th_1, a + th_0) + h_0 \frac{\partial \psi}{\partial y}(a + th_1, a + th_0) \right] dt$$

On vérifie aisément que

$$\frac{\partial \psi}{\partial x}(x, y) = \frac{\tau(x, y) - f'(x)}{\tau(x, y)} + f(x) \frac{\frac{\partial \tau}{\partial x}(x, y)}{\tau(x, y)^2}, \quad \frac{\partial \psi}{\partial y}(x, y) = f(x) \frac{\frac{\partial \tau}{\partial y}(x, y)}{\tau(x, y)^2}$$

Comme $|f(x)| = |f(x) - f(a)| \leq M_1 |x - a|$, les égalités ci-dessus impliquent :

$$\left| \frac{\partial \psi}{\partial x}(x, y) \right| \leq \frac{M_2 |y - x|}{2m_1} + M_1 |x - a| \frac{M_2}{2m_1^2}, \quad \left| \frac{\partial \psi}{\partial y}(x, y) \right| \leq M_1 |x - a| \frac{M_2}{2m_1^2}$$

Pour $(x, y) = (a + th_1, a + th_0)$, on a $|y - x| \leq (|h_0| + |h_1|)t$ et $|x - a| = |h_1|t$ d'où :

$$\begin{aligned} \left| \frac{\partial \psi}{\partial x}(a + th_1, a + th_0) \right| &\leq \left[\frac{M_2}{2m_1} (|h_0| + |h_1|) + \frac{M_1 M_2}{2m_1^2} |h_1| \right] t \\ \left| \frac{\partial \psi}{\partial y}(a + th_1, a + th_0) \right| &\leq \frac{M_1 M_2}{2m_1^2} |h_1| t \\ |h_2| &\leq \left[\frac{M_2}{2m_1} + \frac{M_1 M_2}{2m_1^2} \right] (|h_0| + |h_1|) |h_1| \int_0^1 t dt \\ &= \frac{K}{2} (|h_0| + |h_1|) |h_1| \leq K |h_1| \max(|h_0|, |h_1|) \end{aligned}$$

Comme $|h_0|, |h_1| \leq h \leq \frac{1}{K}$, on voit que $|h_2| \leq |h_1|$. De la même façon, on a aussi :

$$|h_{p+1}| \leq K |h_p| \max(|h_p|, |h_{p-1}|)$$

et ceci entraîne par récurrence que la suite $(|h_p|)_{p \in \mathbb{N}}$ est décroissante. On en déduit

$$|h_{p+1}| \leq K |h_p| |h_{p-1}|$$

L'inégalité $|h_p| \leq \frac{1}{K} [K \max(|h_0|, |h_1|)]^{s_p}$ est évidente pour $p = 0, 1$. Elle résulte de l'estimation précédente pour $p = 2$ et se généralise par récurrence pour $p \geq 3$. ■

4.5 Cas des fonctions vectorielles

On considère à présent une équation $f(x) = 0$ où $f = (f_1, \dots, f_m)$ est une fonction définie sur une partie de \mathbb{R}^m et à valeurs dans \mathbb{R}^m . Comme dans le cas réel, on remplace cette équation par une équation du type $\varphi(x) = x$ où φ est une fonction d'une partie fermée de \mathbb{R}^m dans elle-même, pour pouvoir utiliser le Théorème du point fixe.

On munira l'espace vectoriel \mathbb{R}^m de sa norme euclidienne, c'est à dire $|x| = (x_1^2 + \dots + x_m^2)^{1/2}$ où (x_1, \dots, x_m) sont les coordonnées du vecteur $x \in \mathbb{R}^m$.

On rappelle que les applications linéaires d'un espace vectoriel de dimension finie dans un autre sont continues.

On notera $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ l'espace vectoriel des applications linéaires de \mathbb{R}^m dans lui-même.

La norme d'opérateur d'un élément $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ est définie par :

$$\|T\| = \sup_{x \in \mathbb{R}^m, x \neq 0} \frac{|Tx|}{|x|}.$$

Le rayon spectral d'un élément $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ est défini par :

$$\rho(T) = \max_{1 \leq i \leq m} |\lambda_i|$$

où les λ_i , $1 \leq i \leq m$, sont les valeurs propres de T .

Nous aurons besoin des propriétés du rayon spectral d'un endomorphisme. Pour cela, notons \mathcal{N} l'ensemble des normes sur \mathbb{R}^m et \mathcal{N}_e l'ensemble des normes euclidiennes sur \mathbb{R}^m . La norme d'opérateur de T associée à une norme $N \in \mathcal{N}$ sera notée $\|T\|_N$, c'est à dire :

$$\|T\|_N = \sup_{x \in \mathbb{R}^m, x \neq 0} \frac{N(Tx)}{N(x)}.$$

Rappelons, sans démonstration, le résultat suivant :

Proposition 4.5.1. *Soit T un endomorphisme de \mathbb{R}^m . Alors :*

- 1) *Pour tout $N \in \mathcal{N}$, $\rho(T) \leq \|T\|_N$.*
- 2) *$\rho(T) = \inf_{N \in \mathcal{N}} \|T\|_N = \inf_{N \in \mathcal{N}_e} \|T\|_N$.*
- 3) *Pour tout $N \in \mathcal{N}$, $\rho(T) = \lim_{p \rightarrow +\infty} \|T^p\|_N^{1/p}$.*

On rappelle également que, si φ est une fonction différentiable, définie sur un ouvert $\Omega \subset \mathbb{R}^m$, à valeurs dans \mathbb{R}^m , alors sa différentielle en $x \in \Omega$, notée $D\varphi(x)$, est une application linéaire de \mathbb{R}^m dans \mathbb{R}^m et qu'il existe une fonction ε , définie sur un voisinage de x , à valeurs dans \mathbb{R}^m telle que :

$$\varphi(x+h) - \varphi(x) = D\varphi(x).h + |h|\varepsilon(h), \quad \lim_{h \rightarrow 0} \varepsilon(h) = 0$$

Comme dans le cas réel, on peut définir un cadre où le théorème du point fixe s'applique. Montrons d'abord un résultat un peu plus précis que le Théorème des Accroissements finis pour les fonctions vectorielles :

Lemme 4.5.2. *Soit $\varphi : \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ une fonction de classe C^1 .*

- 1) *Si φ est k -lipschitzienne sur Ω , alors $\|D\varphi(x)\| \leq k$ pour tout $x \in \Omega$.*
- 2) *Si Ω est convexe et si $\|D\varphi(x)\| \leq k$, pour tout $x \in \Omega$, alors φ est k -lipschitzienne sur Ω .*

DÉMONSTRATION : Montrons le 1) :

On écrit :

$$D\varphi(x).h = \varphi(x+h) - \varphi(x) - |h|\varepsilon(h), \quad \lim_{h \rightarrow 0} \varepsilon(h) = 0$$

Pour tout $\delta > 0$, il existe $r > 0$ tel que

$$|h| \leq r \Rightarrow |\varepsilon(h)| \leq \delta$$

D'où, si $|h| \leq r$,

$$|D\varphi(x).h| \leq |\varphi(x+h) - \varphi(x)| + |h| |\varepsilon(h)| \leq (k + \delta) |h|$$

On en déduit $\|D\varphi(x)\| \leq k + \delta$ et ceci pour tout $\delta > 0$, d'où le résultat.

Montrons 2) : Si Ω est convexe, on peut écrire :

$$\varphi(y) - \varphi(x) = \psi(1) - \psi(0) = \int_0^1 \psi'(t) dt, \quad \text{avec } \psi(t) = \varphi(x + t(y-x))$$

Comme $\psi'(t) = D\varphi(x + t(y - x)) \cdot (y - x)$, on en déduit

$$\varphi(y) - \varphi(x) = \int_0^1 D\varphi(x + t(y - x)) \cdot (y - x) dt$$

d'où

$$|\varphi(y) - \varphi(x)| \leq \int_0^1 \|D\varphi(x + t(y - x))\| |y - x| dt \leq k |y - x|.$$

■

Comme pour les fonctions réelles, dans le cas des fonctions sur \mathbb{R}^m de classe C^1 , si φ est contractante sur un voisinage fermé de a , le Théorème du point fixe 4.1.1 permet d'obtenir une convergence de la suite récurrente vers le point fixe. Précisément, on peut énoncer :

Proposition 4.5.3. *Soit f une application continûment différentiable d'un ouvert convexe Ω de \mathbb{R}^m , à valeurs dans \mathbb{R}^m et $a \in \mathbb{R}$ tel que $\varphi(a) = a$. Si $\|D\varphi(a)\| < 1$, alors il existe un voisinage V de a tel que si $x_0 \in V$, la suite récurrente définie par $x_{n+1} = \varphi(x_n)$ converge vers le point fixe a . On dit alors que le point fixe a est attractif.*

DÉMONSTRATION : Puisque la fonction $x \rightarrow D\varphi(x)$ est continue et que $\|D\varphi(a)\| < 1$, il existe $k < 1$ et un voisinage fermé V de a tels que : $\forall x \in V, \|D\varphi(x)\| \leq k$. Puis que Ω est convexe, on peut supposer que V est également convexe. Le Théorème 4.5.2 implique alors que φ est k -lipschitzienne et donc contractante dans V . Puisque $\varphi(a) = a$, V est stable par φ et on peut bien appliquer le Théorème du point fixe, 4.1.1, à la fonction φ sur V . La suite récurrente $x_{n+1} = \varphi(x_n)$ converge donc bien vers le point fixe a . ■

Le résultat suivant donne un critère assurant que la fonction φ est contractante sur un voisinage fermé de a :

Théorème 4.5.4. *Soit φ une fonction de classe C^1 de Ω dans \mathbb{R}^m et soit $a \in \Omega$ un point fixe de φ . Alors, les deux propriétés suivantes sont équivalentes :*

- 1) *Il existe un voisinage fermé V de a tel que $\varphi(V) \subset V$ et une norme N sur \mathbb{R}^m telle que $\varphi|_V$ soit strictement contractante pour N .*
- 2) *$\rho(D\varphi(a)) < 1$, où ρ désigne le rayon spectral.*

DÉMONSTRATION : Si $\varphi|_V$ est contractante pour N , de rapport $k < 1$, alors la proposition 4.5.1 1) montre que :

$$\rho(D\varphi(a)) \leq \|D\varphi(a)\|_N \leq k < 1.$$

Inversement, si $\rho(D\varphi(a)) < 1$, il existe une norme euclidienne N telle que $\|D\varphi(a)\|_N < 1$. Par continuité de $D\varphi$, il existe une boule fermée $V = \bar{B}(a, r)$, $r > 0$, telle que $\sup_{x \in V} \|D\varphi(x)\|_N = k < 1$. Comme V est convexe, par le lemme 4.5.2, φ est alors contractante de rapport k sur V . En particulier, $\varphi(V) \subset \bar{B}(a, kr) \subset V$. ■

Remarque. *Si φ est de classe C^2 et si $D\varphi(a) = 0$, comme dans le cas des fonctions réelles, la formule de Taylor-Young montre qu'il existe une constante $M \geq 0$ telle que*

$$\|\varphi(x) - a\| \leq M \|x - a\|^2$$

Le phénomène de convergence quadratique, déjà observé en dimension 1, a donc encore lieu dans ce cadre.

Il existe une extension de la méthode de Newton pour approcher les solutions a d'équations vectorielles $f(x) = 0$ à l'aide du Théorème du point fixe, appelée méthode de Newton-Raphson, qui consiste encore à remplacer la fonction f par sa partie linéaire en un point voisin x_0 , c'est à dire par

$$f(x_0) + Df(x_0)(x - x_0)$$

On résout alors l'équation $f(x_0) + Df(x_0)(x - x_0) = 0$. Si $Df(x_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ est inversible, on a une solution unique x_1 telle que $x_1 - x_0 = -Df(x_0)^{-1}f(x_0)$.

On va donc itérer l'application de classe C^1 :

$$\varphi(x) = x - Df(x)^{-1}f(x)$$

Théorème 4.5.5. (Méthode de Newton-Raphson) $f : \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ une fonction de classe C^2 et $a \in \Omega$, tel que $f(a) = 0$. Supposons que l'application linéaire tangente à f en a , $Df(a) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ soit inversible. Alors, a est un point super attractif de la fonction φ définie par

$$\varphi(x) = x - Df(x)^{-1}.f(x).$$

DÉMONSTRATION : On notera $D^2f(a)$ la différentielle seconde de f en a . On rappelle que $D^2f(a)$ est une forme bilinéaire symétrique sur \mathbb{R}^m . En vue de trouver un développement de Taylor de φ en a à l'ordre 2, écrivons les développements de Taylor de f et Df en a :

$$\begin{aligned} f(a+h) &= Df(a).h + \frac{1}{2}D^2f(a)(h, h) + o(|h|^2) \\ &= Df(a).[h + \frac{1}{2}Df(a)^{-1}.D^2f(a)(h, h) + o(|h|^2)] \end{aligned}$$

$$Df(a+h) = Df(a) + D^2f(a).h + o(|h|) = Df(a)[Id + Df(a)^{-1}.D^2f(a).h + o(|h|)]$$

Par continuité de Df , il existe un voisinage de a sur lequel $Df(x)$ est encore inversible et donc si h est suffisamment petit, on peut écrire :

$$\begin{aligned} Df(a+h)^{-1} &= [Id + Df(a)^{-1}.D^2f(a).h + o(|h|)]^{-1}Df(a)^{-1} \\ &= [Id - Df(a)^{-1}.D^2f(a).h + o(|h|)]Df(a)^{-1} \end{aligned}$$

D'où

$$\begin{aligned} Df(a+h)^{-1}f(a+h) &= \\ [Id - Df(a)^{-1}.D^2f(a).h + o(|h|)] [h + \frac{1}{2}Df(a)^{-1}.D^2f(a)(h, h) + o(|h|^2)] &= \\ = h - \frac{1}{2}Df(a)^{-1}.D^2f(a).(h, h) + o(|h|^2) \end{aligned}$$

Finalement on obtient

$$\varphi(a+h) = a+h - Df(a+h)^{-1}f(a+h) = a + \frac{1}{2}Df(a)^{-1}.D^2f(a).(h, h) + o(|h|^2)$$

On en déduit $D\varphi(a) = 0$ et $D^2\varphi(a) = Df(a)^{-1}.D^2f(a)$. En particulier :

$$|\varphi(a+h) - a| \leq \frac{1}{2}(M + \varepsilon(h)) |h|^2$$

où $M = \|D^2\varphi(a)\|$. ■

Exemple 4.5.6. Résoudre le système

$$\begin{cases} x^2 + xy - 2y^2 = 4 \\ xe^x + ye^y = 0 \end{cases}.$$

La courbe $C_1 : x^2 + xy - 2y^2 = 4$ est une hyperbole d'asymptotes $x^2 + xy - 2y^2 = (x - y)(x + 2y) = 0$ qui passe par les points $(2, 0)$ et $(-2, 0)$.

La courbe $C_2 : xe^x + ye^y = 0$ est symétrique par rapport à la droite $y = x$. De plus, x et y sont nécessairement de signes opposés. Supposons par exemple $x \leq 0$ et $y \geq 0$. Comme la fonction $y \rightarrow ye^y$ est strictement croissante de $[0, +\infty[$ sur lui-même, à chaque point $x \leq 0$ correspond un unique point $y \geq 0$, solution de $xe^x + ye^y = 0$. On peut par exemple obtenir une approximation de ce point par itération de la fonction

$$\varphi(x) = y - \frac{xe^x + ye^y}{(1 + y)e^y} = \frac{y^2 - xe^{x-y}}{1 + y}$$

fournie par la méthode de Newton appliquée à la variable y .

En traçant les courbes C_1 et C_2 , on voit que le système admet une solution unique (a, b) voisine de $(-2; 0, 2)$. Pour obtenir une approximation plus précise de cette solution, on cherche à résoudre l'équation $f(x, y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ avec

$$f(x, y) = \begin{pmatrix} x^2 + xy - 2y^2 - 4 \\ xe^x + ye^y \end{pmatrix}$$

L'application linéaire tangente à f en (x, y) est donnée par la matrice :

$$Df(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x, y) & \frac{\partial f_1}{\partial y}(x, y) \\ \frac{\partial f_2}{\partial x}(x, y) & \frac{\partial f_2}{\partial y}(x, y) \end{pmatrix} = \begin{pmatrix} 2x + y & x - 4y \\ (x + 1)e^x & (y + 1)e^y \end{pmatrix}$$

La condition $Df(a, b)$ inversible signifie que les tangentes

$$\frac{\partial f_i}{\partial x}(a, b)(x - a) + \frac{\partial f_i}{\partial y}(a, b)(y - b) = 0, \quad i = 1, 2$$

aux courbes C_1 et C_2 au point (a, b) sont distinctes. C'est bien le cas ici. On obtient :

$$\left[f'(x, y) \right]^{-1} = \frac{1}{\Delta(x, y)} \begin{pmatrix} (y + 1)e^y & -x + 4y \\ -(x + 1)e^x & 2x + y \end{pmatrix}$$

où $\Delta(x, y) = (2x + y)(y + 1)e^y - (x - 4y)(x + 1)e^x$.

On est alors amené à calculer les itérés $(x_{p+1}, y_{p+1}) = \varphi(x_p, y_p)$ avec :

$$\varphi(x, y) = \begin{pmatrix} x \\ y \end{pmatrix} - \frac{1}{\Delta(x, y)} \begin{pmatrix} (y + 1)e^y & -x + 4y \\ -(x + 1)e^x & 2x + y \end{pmatrix} \begin{pmatrix} x^2 + xy - 2y^2 - 4 \\ xe^x + ye^y \end{pmatrix}$$

en partant du point initial $(-2; 0, 2)$. On trouve par exemple :

$$(a, b) \sim (x_4, y_4) = (-2, 126932304; 0, 206278156)$$

5

EQUATIONS DIFFÉRENTIELLES, EXISTENCE ET UNICITÉ DES SOLUTIONS

5.1 Introduction

Comme au chapitre 4, on note \mathbb{R}^m l'espace Euclidien à m dimensions, et on note $|x|$ la norme euclidienne d'un point $x = (x_1, \dots, x_m)$, définie par $|x|^2 = x_1^2 + \dots + x_m^2$ et $\langle x, y \rangle = x_1 y_1 + \dots + x_m y_m$ le produit scalaire des vecteurs $x = (x_1, \dots, x_m)$ et $y = (y_1, \dots, y_m)$. On se donne un intervalle compact d'intérieur non vide $I \subset \mathbb{R}$, un ouvert non vide $V \subset \mathbb{R}^m$. Le plus souvent on prendra $I = [0, T]$ avec $T > 0$ et $V = \mathbb{R}^m$. On se donne une application continue

$$(t, x) \in I \times V \rightarrow f(t, x) \in \mathbb{R}^m.$$

Définition 5.1.1. On appelle *solution* (solution globale s'il y a ambiguïté) de l'équation différentielle ordinaire EDO

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

toute fonction continue $t \in I \rightarrow y(t) \in \mathbb{R}^m$, prenant ses valeurs dans l'ouvert V et telle que, en tout point t de l'intérieur de I , y est dérivable et la dérivée $y'(t)$ vérifie

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

On peut étendre cette définition aux intervalles I quelconques de \mathbb{R} , en particulier $I = [0, +\infty[$:

Définition 5.1.2. Sur un intervalle I quelconque de \mathbb{R} , on appelle *solution* (solution globale s'il y a ambiguïté) de l'équation différentielle ordinaire EDO

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

toute fonction continue $y : I \rightarrow \mathbb{R}^m$ qui est solution de l'EDO sur tout sous-intervalle compact $J \subset I$.

Notation 5.1.3. On appellera toujours temps la variable t et, parfois, espace de configuration l'ouvert V , dont la dimension $m \geq 1$ est le nombre de degrés de liberté.

C'est un vocabulaire héritée de la mécanique, science qui est à l'origine du calcul différentiel.

Notons que *EDO* peut aussi s'écrire avec les coordonnées des vecteurs sous la forme d'un système de m équations scalaires :

$$\begin{aligned} y_1'(t) &= f_1(t, y_1(t), \dots, y_m(t)) \\ &\dots \\ y_m'(t) &= f_m(t, y_1(t), \dots, y_m(t)) \end{aligned}$$

où les $y_k(t)$ sont des fonctions à valeurs réelles, mais c'est beaucoup plus lourd.

Définition 5.1.4. *Si, étant donné un temps $t_0 \in I$ et un point $y_0 \in V$, on impose de plus à une solution y de EDO la condition :*

$$y(t_0) = y_0,$$

on parlera de solution passant par y_0 à l'instant t_0 et le couple $(t_0, y_0) \in I \times \mathbb{R}^m$ sera appelé condition initiale de l'équation. On dira dans ce cas, que l'on cherche à résoudre l'équation différentielle EDO avec la condition initiale $y(t_0) = y_0$.

En général, on se limitera au cas $I = [0, T]$, où $T > 0$ est un temps final fixé, et parfois au cas limite $I = [0, +\infty[$, correspondant à un temps final infini. Alors, la condition initiale sera, dans ce cas, donnée au temps $t_0 = 0$.

Se limiter au cas $t_0 = 0$ n'est pas vraiment un cas restrictif, en vertu des résultats élémentaires suivants qui permettent de s'y ramener par changement de variables :

Proposition 5.1.5. *Soit deux fonctions continues y et f à valeurs dans \mathbb{R}^m définies respectivement sur $[0, T]$ et $[0, T] \times V$. Soit t_0 réel quelconque fixé. Posons $I = [t_0, t_0 + T]$ et*

$$g(t, x) = f(t - t_0, x), \quad z(t) = y(t - t_0),$$

pour tous $t \in I$ et $x \in V$. Alors z est solution sur I de l'EDO $z'(t) = g(t, z(t))$ si et seulement si y est solution de l'EDO $y'(u) = f(u, y(u))$ sur l'intervalle $[0, T]$.

Proposition 5.1.6. *Soit deux fonctions continues y et f à valeurs dans \mathbb{R}^m définies respectivement sur $[0, T]$ et $[0, T] \times V$. Posons $J = [-T, 0]$ et*

$$g(t, x) = -f(-t, x), \quad z(t) = y(-t),$$

pour tous $t \in I$ et $x \in V$. Alors z est solution sur J de l'EDO $z'(t) = g(t, z(t))$ si et seulement si y est solution de l'EDO $y'(s) = f(s, y(s))$ sur l'intervalle $[0, T]$.

DÉMONSTRATION :

Dans la première proposition, il suffit de faire le changement de variable $u = t - t_0$ et dans la seconde $s = -t$. ■

Souvent, l'ouvert V sera \mathbb{R}^m tout entier, mais il y a des applications importantes où cela ne sera pas le cas, par exemple, dans le problème problèmes à N corps que nous étudierons plus loin.

Définition 5.1.7. 1) Lorsque la fonction f ne dépend pas de t , on dira que l'équation EDO est autonome.

2) Si $f(t, x)$ dépend affinement de x , on dira que l'équation est linéaire. (On ne demande pas de dépendance linéaire ou affine par rapport à t .)

Exemples 5.1.8. Cas d'un degré de liberté.

Le cas le plus simple est celui où $m = 1$ et $V = \mathbb{R}$; l'équation linéaire la plus simple est

$$y'(t) = ay(t)$$

où a est un réel fixé, ce qui correspond à la fonction d'une variable réelle $f(x) = ax$.

La solution s'écrit :

$$y(t) = y(0) \exp(at)$$

Cette équation est un modèle élémentaire de croissance si $a > 0$, ou d'extinction si $a < 0$, qu'on retrouve un peu partout en démographie, économie, finance, biologie etc.... On appelle taux de croissance le paramètre a .

L'équation linéaire générale est

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

où a et b sont des fonctions données. Elle a pour solution

$$y(t) = y(0) \exp\left(\int_0^t a(s) ds\right) + \int_0^t b(\theta) \exp\left(\int_\theta^t a(s) ds\right) d\theta$$

On retrouvera ce résultat au chapitre 6 par la fameuse méthode de variation de la constante. Bien sur, en termes de modèle de croissance $a(t)$ est un taux variable et $b(t)$ décrit des apports, positifs ou négatifs, externes.

L'exemple le plus simple d'équation non-linéaire

$$y'(t) = y^2(t)$$

dont la solution vaut

$$y(t) = \frac{y(0)}{1 - ty(0)}$$

fournit un exemple de solution seulement locale, sur $I = [0, +\infty[$, dès que $y(0) > 0$, puisqu'alors $y(t) \rightarrow +\infty$ lorsque t s'approche de $y(0)^{-1}$.

Cette équation est un modèle de croissance explosive.

Toute équation ne peut pas être résolue aussi simplement par quadrature, c'est-à-dire à l'aide de fonctions algébriques et spéciales, telles que l'exponentielle ou les fonctions trigonométriques... et de leurs primitives. Joseph Liouville a fourni le premier contre-exemple avec l'équation non-linéaire non-autonome, mais pourtant fort simple,

$$y'(t) = y^2(t) - t$$

Exemple 5.1.9. Cas de deux degrés de liberté.

Avec deux degrés de liberté, c'est-à-dire $m = 2$, on a de nombreux exemples tirés de la mécanique, qui sont en fait des équations différentielles du deuxième ordre

$$\varphi''(t) = f(t, \varphi(t), \varphi'(t))$$

On se ramène à une équation du premier ordre à deux degrés de liberté grâce à la proposition suivante :

Proposition 5.1.10. *Toute équation différentielle du deuxième ordre*

$$\varphi''(t) = f(t, \varphi'(t), \varphi(t))$$

est équivalente à l'équation différentielle du premier ordre sur \mathbb{R}^2 :

$$y_1'(t) = y_2(t), \quad y_2'(t) = f(t, y_1(t), y_2(t)).$$

DÉMONSTRATION : Pour faire la correspondance, il suffit de considérer la fonction vectorielle $Y = (y_1, y_2)$ à valeurs dans \mathbb{R}^2 telle que :

$$y_1 = \varphi, \quad y_2 = \varphi'$$

ce qui revient à éclater la dérivée seconde en deux dérivées premières.

L'équation du premier ordre obtenue s'écrit alors sous forme vectorielle :

$$Y'(t) = \begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} y_2(t) \\ f(t, y_1(t), y_2(t)) \end{pmatrix}$$

■

Remarque 5.1.11. *La condition initiale définie pour une équation du premier ordre à valeurs dans \mathbb{R}^2 , c'est à dire $Y(t_0) = \begin{pmatrix} y_1(t_0) \\ y_2(t_0) \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2$ correspond, dans le cas où cette équation provient d'une équation du deuxième ordre à valeurs dans \mathbb{R} à :*

$$\varphi(t_0) = a_1, \quad \varphi'(t_0) = a_2$$

Ainsi la dynamique d'un pendule oscillant autour de l'axe vertical sous l'effet de la gravité, est décrite par sa longueur L et son angle φ par rapport à la verticale au moyen de l'équation différentielle du second ordre

$$\varphi'' + \frac{g}{L} \sin \varphi = 0$$

où $g > 0$ est l'accélération de la pesanteur. On ramène cette équation dans le cadre général avec $m = 2$, $V = \mathbb{R}^2$, en introduisant

$$f(y) = (y_2, -\frac{g}{L} \sin(y_1))$$

avec :

$$y_1' = y_2, \quad y_2' = -\frac{g}{L} \sin(y_1).$$

Un exemple plus simple que le pendule est l'équation linéarisée du pendule

$$\varphi'' + \frac{g}{L} \varphi = 0$$

qui est une approximation de la précédente, d'autant plus acceptable que l'angle φ autour de la verticale reste faible.

Sa solution générale est bien connue

$$\varphi(t) = \varphi(0) \cos(\omega t) + \omega^{-1} \varphi'(0) \sin(\omega t)$$

où $\omega = (\frac{g}{L})^{1/2}$.

Elle vous permet de mesurer g assez facilement avec un bout de ficelle.

Un autre exemple est celui de l'équation décrivant, en première approximation, l'orbite d'une planète autour du soleil. Si on note $r(t) > 0$ la distance au soleil, l'équation est donnée par

$$r'' = -k/r^2 + M^2/r^3$$

où k et M sont des paramètres physiques. On se ramène à la forme standard d'ordre 1 en posant $m = 2$, $V =]0, +\infty[\times \mathbb{R}$ et

$$f(y) = (y_2, -k/y_1^2 + M^2/y_1^3).$$

C'est le premier de nos exemples où V n'est pas \mathbb{R}^m tout entier. Cette équation s'intègre complètement et permet de retrouver les fameuses lois de Kepler régissant le mouvement des planètes.

Exemple 5.1.12. *Cas de m degrés de liberté, $m > 2$.*

Comme pour le cas de deux degrés de liberté, on trouve fréquemment des équations d'ordre m sur \mathbb{R} et on se ramène à des équations d'ordre 1 à m degrés de liberté :

Proposition 5.1.13. *Toute équation différentielle d'ordre m sur \mathbb{R} :*

$$\varphi^{(m)}(t) = f(t, \varphi(t), \varphi'(t), \dots, \varphi^{(m-1)}(t))$$

est équivalente à l'équation différentielle du premier ordre sur \mathbb{R}^m :

$$y_1'(t) = y_2(t), \quad y_2'(t) = y_3(t), \dots, y_m'(t) = f(t, y_1(t), \dots, y_m(t)).$$

DÉMONSTRATION : Comme pour $m = 2$, on considère la fonction vectorielle $Y = (y_1, \dots, y_m)$ à valeurs dans \mathbb{R}^m telle que :

$$y_1 = \varphi, \quad y_2 = \varphi', \dots, y_m = \varphi^{(m-1)}$$

ce qui revient à éclater la dérivée m -ième en m dérivées premières.

L'équation du premier ordre obtenue s'écrit alors sous forme vectorielle :

$$Y'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_m'(t) \end{pmatrix} = \begin{pmatrix} y_2(t) \\ \vdots \\ f(t, y_1(t), \dots, y_m(t)) \end{pmatrix}$$

■

Remarque 5.1.14. La condition initiale définie pour une équation du premier ordre à

valeurs dans \mathbb{R}^m , c'est à dire $Y(t_0) = \begin{pmatrix} y_1(t_0) \\ \vdots \\ y_m(t_0) \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{R}^m$ correspond, dans le cas

où cette équation provient d'une équation d'ordre m à valeurs dans \mathbb{R} à :

$$\varphi(t_0) = a_1, \varphi'(t_0) = a_2, \dots, \varphi^{(m-1)}(t_0) = a_m$$

Remarque. Selon le même principe, une équation différentielle d'ordre k sur \mathbb{R}^m se ramène à une équation d'ordre 1 sur \mathbb{R}^{km} , avec la condition initiale correspondante.

L'exemple le plus célèbre est celui des N corps qui décrit l'interaction de N corps célestes en interaction suivant la loi d'attraction de Newton, par exemple le soleil et les 9 planètes. On note $x_\alpha(t) \in \mathbb{R}^3$ la position dans l'espace au temps t du corps numéro α , α allant de 1 à N , et m_α sa masse. Pour simplifier, on assimile ces corps à des points. Les équations de Newton sont

$$x''_\alpha = -G \sum_{\beta \neq \alpha} m_\beta \frac{x_\alpha - x_\beta}{|x_\alpha - x_\beta|^3}, \quad \alpha = 1, \dots, N$$

et sont valables tant qu'il n'y a pas de collisions.

On peut ramener ce système de $3N$ équations couplées à une équation autonome de la forme standard sur un espace de configuration de dimension $m = 6N$, le 6 étant dû à l'éclatement des dérivées secondes comme dans le cas précédent, soit 60 pour le système solaire. On posera

$$x_1(t) = (y_1(t), y_2(t), y_3(t)), \dots, x_N(t) = (y_{3N-2}(t), y_{3N-1}(t), y_{3N}(t)),$$

$$x'_1(t) = (y_{3N+1}(t), y_{3N+2}(t), y_{3N+3}(t)), \dots, x'_N(t) = (y_{6N-2}(t), y_{6N-1}(t), y_{6N}(t)),$$

On notera que la fonction f correspondante ne peut être défini en tous points de \mathbb{R}^{6N} , mais seulement sur l'ouvert V des points y de \mathbb{R}^{6N} dont les N premiers paquets de 3 coordonnées sont deux à deux différents, ce qui revient à interdire les collisions.

On peut donc s'attendre à obtenir parfois des solutions seulement locales, des collisions se produisant au bout d'un temps fini : il se peut aussi que des corps partent à l'infini en temps fini.

La résolution de ce système ne peut s'effectuer par quadrature que pour $N = 2$, problème résolu par Newton qui a ainsi retrouvé les lois de Kepler. On se ramène à ce cas en faisant l'approximation que la masse de toutes les planètes est négligeable par rapport à celle du soleil. On pose donc $m_\alpha = 0$ sauf pour $\alpha = 1$. Le système se découple alors complètement.

Exemple d'Approximation 5.1.15. Le schéma d'Euler.

Bien avant l'informatique et le calcul scientifique, il est apparu judicieux d'introduire des procédures de calcul approché pour les EDO. L'idée la plus simple est de discrétiser le temps. Plaçons-nous sur l'intervalle $I = [0, +\infty[$ et supposons qu'on veuille approcher une solution $y(t)$ définie pour $t \geq 0$. On introduit un laps de temps $h > 0$, appelé pas de temps, et on considère les instants multiples entiers de h , $t = nh$ pour $n = 0, 1, 2, \dots$. A chacun de ces instants, on a

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

Mais, dans la mesure où y est une fonction suffisamment régulière, $y'(nh)$ est peu différent du quotient différentiel

$$\frac{y((n+1)h) - y(nh)}{h}.$$

Il est donc naturel d'introduire la suite numérique y_n définie par récurrence par

$$\frac{y_{n+1} - y_n}{h} = f(nh, y_n), \quad n = 0, 1, 2, \dots$$

avec $y_0 = y(0)$ pour démarrer.

Cette suite dépend de h , on peut la noter plus lourdement $y_{n,h}$ avec $y_{0,h} = y(0)$. Elle permet de construire une fonction $y_h(t)$ affine par morceaux définie pour $t \in I = [0, +\infty[$ et obtenue en interpolant les y_n aux instants $t = nh$, c'est à dire :

$$y_h(t) = y_n + (t - nh)f(nh, y_n) \text{ pour } nh \leq t \leq (n+1)h$$

Une façon équivalente de définir cette fonction est de noter pour $t \geq 0$, h fois la partie entière de t/h par t_h , et de poser,

$$y_h(t) = y_h(t_h) + (t - t_h)f(t_h, y_h(t_h)).$$

On s'attend à ce que y_h soit une approximation de la solution y , convergente lorsque le pas de temps h tend vers 0. Pour mesurer cette convergence, on évaluera l'erreur au n -ième pas de temps, égale à la différence entre la valeur prise par la solution y au point nh et celle prise par l'approximation y_h , c'est à dire

$$e_n = y_h(nh) - y(nh) = y_n - y(nh)$$

Dans tous les cas, au moins lorsque f est définie sur $I \times \mathbb{R}^m$, la solution approchée est calculable numériquement à l'aide de simples évaluations de la fonction f .

Exemple 5.1.16. *Considérons l'équation $y' = ay$.*

On remarquera que, pour $t > 0$ fixé, en choisissant $h = t/n$, avec $n > 0$ entier, la solution approchée $y_h(t)$ est tout simplement donnée par

$$y_h(t) = (1 + atn^{-1})^n y(0).$$

La convergence de $y_h(t)$ vers $y(t) = y(0) \exp(at)$ quand $h = t/n$ tend vers 0 est alors assurée par le résultat bien connu d'analyse

$$e^b = \lim_{n \rightarrow \infty} (1 + b/n)^n$$

pour tout réel b fixé.

Dans le chapitre 7 de ce cours on établira la convergence du schéma d'Euler sous des hypothèses convenables sur la donnée f .

5.2 Principaux résultats dans le cas Lipschitzien

Le cadre le plus simple pour aborder les questions théoriques est celui d'une équation avec $V = \mathbb{R}^m$, où

$$(t, x) \in [0, T] \times \mathbb{R}^m \rightarrow f(t, x) \in \mathbb{R}^m$$

est une fonction continue définie partout sur $[0, T] \times \mathbb{R}^m$.

Commençons par une définition :

Définition 5.2.1. On dit qu'une fonction g d'un espace vectoriel E muni d'une norme $\|\cdot\|_E$ à valeurs dans un espace vectoriel F muni d'une norme $\|\cdot\|_F$ est Lipschitzienne s'il existe une constante L telle que

$$\|g(c) - g(b)\|_F \leq L\|c - b\|_E, \quad \forall b, c \in E.$$

La meilleure constante L est appelée constante de Lipschitz et notée $Lip(g)$.

Dans le cas particulier où cette constante est strictement inférieure à 1, avec $E = F$ et $\|\cdot\|_E = \|\cdot\|_F$, comme au chapitre précédent, on dit que l'application est contractante.

Il est clair que si g est Lipschitzienne, alors cette fonction est uniformément continue.

Si $E = \mathbb{R}^m$ et $F = \mathbb{R}^p$ et si g est continument différentiable, une condition nécessaire et suffisante pour que g soit Lipschitzienne est que les dérivées partielles $\frac{\partial g_j(x)}{\partial x_i}$, pour $i = 1, \dots, m, j = 1, \dots, p$ soient bornées.

Exemples 5.2.2. 1) Comme exemples de fonctions Lipschitziennes de \mathbb{R}^m dans \mathbb{R} , donnons $g(x) = |x|, g(x) = \sin(|x|)$

2) Comme exemples de fonctions continues mais pas Lipschitziennes, donnons $g(x) = |x|^\alpha$ pour $\alpha > 0$ différent de 1.

3) Comme exemple fondamental d'application lipschitzienne de \mathbb{R}^m dans lui-même, citons $g(x) = A.x$ où A est une matrice carrée $m \times m$. La constante de Lipschitz de g est alors la norme matricielle de $A, \|A\|$, associée à la norme euclidienne sur \mathbb{R}^m .

Dans le cadre des équations différentielles, on va faire sur la donnée $f(t, x)$ de l'EDO l'hypothèse d'être Lipschitzienne relativement à la variable x , uniformément en t :

Définition 5.2.3. On dit qu'une fonction $f(t, x)$, définie et continue de $[0, T] \times \mathbb{R}^m$ dans \mathbb{R}^m est uniformément Lipschitzienne relativement à la variable x , uniformément en t s'il existe une constante L telle que

$$|f(t, c) - f(t, b)| \leq L|c - b|, \quad \forall b, c \in \mathbb{R}^m, \quad \forall t \in [0, T].$$

Cette hypothèse n'est pas toujours satisfaite dans les applications. Parmi les exemples qu'on a vus, les EDO : $y'(t) = ay(t), y'(t) = a(t)y + b(t)$ avec $a(t)$ bornée et $\varphi'' + gL \sin \varphi = 0$ vérifient cette hypothèse les autres ne la vérifient pas.

Théorème 5.2.4. (Théorème de Cauchy-Lipschitz) Soit $T > 0$ fixé, a donné dans \mathbb{R}^m et f fonction continue donnée de $[0, T] \times \mathbb{R}^m$ dans \mathbb{R}^m et satisfaisant 5.2.3. Alors il existe une unique fonction continue

$$t \in [0, T] \rightarrow y(t) \in \mathbb{R}^m$$

solution de l'EDO $y'(t) = f(t, y(t))$ avec la condition initiale $y(0) = a$.

Avant de démontrer le Théorème de Cauchy-Lipschitz, on peut en déduire l'existence et l'unicité des solutions d'équations d'ordre m sur \mathbb{R} , conformément à la Proposition 5.1.13 :

Corollaire 5.2.5. Soit $T > 0$ fixé, a donné dans \mathbb{R}^m et f fonction continue de $[0, T] \times \mathbb{R}^m$ dans \mathbb{R} et satisfaisant 5.2.3. Alors il existe une unique fonction continue

$$t \in [0, T] \rightarrow \varphi(t) \in \mathbb{R}$$

solution de l'EDO $\varphi^{(m)}(t) = f(t, \varphi(t), \varphi'(t), \dots, \varphi^{(m-1)}(t))$ avec la condition initiale $\varphi(t_0) = a_1$, $\varphi'(t_0) = a_2$, ..., $\varphi^{(m-1)}(t_0) = a_m$.

DÉMONSTRATION : On considère la fonction vectorielle $Y = (y_1, \dots, y_m)$ à valeurs dans \mathbb{R}^m telle que :

$$y_1 = \varphi, \quad y_2 = \varphi', \quad \dots, \quad y_m = \varphi^{(m-1)}$$

L'équation du premier ordre obtenue s'écrit alors sous forme vectorielle :

$$Y'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_m'(t) \end{pmatrix} = \begin{pmatrix} y_2(t) \\ \vdots \\ f(t, y_1(t), \dots, y_m(t)) \end{pmatrix}$$

avec la condition initiale définie $Y(t_0) = \begin{pmatrix} y_1(t_0) \\ \vdots \\ y_m(t_0) \end{pmatrix} = \begin{pmatrix} \varphi(t_0) \\ \vdots \\ \varphi^{(m-1)}(t_0) \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{R}^m$.

On peut donc appliquer le Théorème de Cauchy-Lipschitz en vérifiant que la fonction $F : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, définie par

$$F(t, x_1, \dots, x_m) = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ f(t, x_1, x_2, \dots, x_m) \end{pmatrix}$$

est bien uniformément Lipschitzienne par rapport à $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ uniformément en t , ce qui

est évident par hypothèse. ■

La preuve du théorème de Cauchy-Lipschitz, 5.2.4, utilise les propositions suivantes :

Proposition 5.2.6. Soit y une fonction continue sur $[0, T]$ à valeurs dans \mathbb{R}^m . Alors, y est solution de l'EDO

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

avec la condition initiale $y(0) = a$ si et seulement si

$$y(t) = a + \int_0^t f(s, y(s)) ds, \quad \forall t \in [0, T].$$

DÉMONSTRATION : On observe que dès qu'une fonction continue y vérifie $y(t) = a + \int_0^t f(s, y(s)) ds$, $\forall t \in [0, T]$, elle est automatiquement dérivable avec une dérivée continue. On peut alors dériver par rapport à t , ce qui donne bien $y'(t) = f(t, y(t))$.

Réciproquement, on obtient le résultat par intégration en t sachant que $y(0) = a$. ■

Proposition 5.2.7. (Lemme de Gronwall, version différentielle) Soit deux réels $L \neq 0$ et b . Soit une fonction continue z de $[0, T]$ dans \mathbb{R} , telle qu'en tout point $t \in]0, T[$, z soit dérivable et

$$z'(t) \leq Lz(t) + b$$

pour tout $t \in]0, T[$. Alors

$$z(t) + b/L \leq (z(0) + b/L) \exp(Lt), \quad \forall t \in [0, T].$$

DÉMONSTRATION : L'idée est de faire un changement de fonction en écrivant z sous la forme

$$z(t) = \zeta(t) \exp(Lt).$$

L'inégalité voulue est alors strictement équivalente à

$$\zeta'(t) \leq b \exp(-Lt).$$

En intégrant de 0 à t , on déduit

$$\zeta(t) - \zeta(0) \leq (1 - \exp(-Lt))b/L,$$

ce qui donne facilement la solution. ■

Proposition 5.2.8. (Lemme de Gronwall, version discrète) Soit deux réels $r \neq 0$ et b . Soit une suite de réels z_n , $n = 0, 1, 2, \dots$ telle que

$$z_{n+1} \leq (1 + r)z_n + b$$

pour tout $n = 0, 1, 2, \dots$. Alors

$$z_n + b/r \leq (z_0 + b/r)(1 + r)^n \leq (z_0 + b/r) \exp(nr)$$

pour tout $n = 0, 1, 2, \dots$

Rappel 5.2.9. Soit l'espace des fonctions continues sur $[0, T]$ à valeurs dans l'espace euclidien \mathbb{R}^m , muni de la norme uniforme

$$\|y\| = \sup_{t \in [0, T]} |y(t)|.$$

Cet espace, noté $\mathcal{C}([0, T], \mathbb{R}^m)$, est un espace de Banach, c'est-à-dire un espace vectoriel normé complet. La complétude signifie que toute suite de Cauchy converge. Autrement dit, pour toute suite $(y_n)_{n \in \mathbb{N}}$ de fonctions continues de $[0, T]$ dans \mathbb{R}^m qui vérifie le critère

de Cauchy pour la norme uniforme, il existe une fonction continue y vers laquelle la suite converge uniformément.

Remarque 5.2.10. Une condition suffisante pour que la suite $(y_n)_{n \in \mathbb{N}}$ vérifie le critère de Cauchy dans $\mathcal{C}([0, T], \mathbb{R}^m)$ est la convergence de la série numérique, à termes positifs :

$$\sum_n \|y_{n+1} - y_n\|.$$

DÉMONSTRATION : En effet on déduit de l'inégalité du triangle que pour tous $p < q$ entiers plus grand que n

$$\|y_p - y_q\| = \|y_p - y_{p+1} + \dots + y_{q-1} - y_q\| \leq \sum_{p \leq k < q} \|y_{k+1} - y_k\| \leq \sum_{p \leq k} \|y_{k+1} - y_k\|,$$

Si la série $\sum_n \|y_{n+1} - y_n\|$ converge, son reste tend vers 0 quand n tend vers l'infini et l'inégalité précédente implique bien qu'alors la suite $(y_n)_{n \in \mathbb{N}}$ vérifie le critère de Cauchy dans $\mathcal{C}([0, T], \mathbb{R}^m)$.

Quand l'espace est complet, on sait alors que la suite $(y_n)_{n \in \mathbb{N}}$ converge vers une limite y et on obtient alors, en posant $p = n$ et en faisant tendre q vers l'infini, l'estimation d'erreur entre y_n et la limite y :

$$\|y_n - y\| \leq \sum_{k \geq n} \|y_{k+1} - y_k\|.$$

■

DÉMONSTRATION : (du théorème 5.2.4) : Unicité de la solution

Supposons qu'on ait deux solutions y et \tilde{y} . Elles sont continument différentiables et on peut donc calculer la dérivée par rapport à t de $|y(t) - \tilde{y}(t)|^2$. Comme $|\cdot|$ est la norme Euclidienne, on obtient, à partir de l'expression : $y'(t) = f(t, y(t))$,

$$\frac{d}{dt} |y(t) - \tilde{y}(t)|^2 = 2 \langle y(t) - \tilde{y}(t), f(t, y(t)) - f(t, \tilde{y}(t)) \rangle,$$

En utilisant l'inégalité de Cauchy-Schwarz et la propriété de Lipschitz (5.2.3), on trouve l'inégalité différentielle

$$\frac{d}{dt} |y(t) - \tilde{y}(t)|^2 \leq 2L |y(t) - \tilde{y}(t)|^2.$$

Grace à la proposition 5.2.7, on trouve, pour tous $t \geq 0$,

$$|y(t) - \tilde{y}(t)|^2 \leq |y(0) - \tilde{y}(0)|^2 \exp(2Lt)$$

et, après avoir pris les racines carrées,

$$|y(t) - \tilde{y}(t)| \leq |y(0) - \tilde{y}(0)| \exp(Lt).$$

Ceci nous donne l'unicité de la solution pour une condition initiale a , puisqu'alors $y(0) = \tilde{y}(0) = a$ et donc $y(t) = \tilde{y}(t)$ a lieu pour tous $t \geq 0$. Cela nous donne en prime un résultat important sur la dépendance Lipschitzienne des solutions au temps t par rapport aux valeurs initiales avec une constante de Lipschitz bornée par $\exp(Lt)$.

Une petite erreur sur les valeurs initiales conduit à une petite erreur au temps t mais avec une amplification exponentielle par rapport à t : c'est ce qu'il est convenu d'appeler l'effet papillon.

Remarque. La démonstration est inchangée si f satisfait la condition plus faible

$$(f(t, c) - f(t, b)) \cdot (c - b) \leq L|c - b|^2, \quad \forall b, c \in E, \quad \forall t \in [0, T].$$

Prouvons maintenant l'existence de solutions :

On peut montrer la fin du théorème 5.2.4, c'est-à-dire l'existence d'une solution, directement en prouvant la convergence de la méthode d'Euler, ce qui constitue l'approche la plus constructive. C'est ce que l'on fera en appendice dans le cas autonome. Mais, pour l'instant, on va se limiter au seul résultat d'existence, en utilisant une preuve beaucoup plus simple mais moins utilisable en pratique pour le calcul numérique, par la méthode des approximations successives ou méthode de Picard.

On définit une suite de fonctions continues $(y_n)_{n \in \mathbb{N}}$, en posant

$$y_{n+1}(t) = a + \int_0^t f(s, y_n(s)) ds, \quad \forall t \in [0, T]$$

avec $y_0(t) = a$ pour tout $t \in [0, T]$.

Par différence des itérations n et $n + 1$, on obtient

$$y_{n+1}(t) - y_n(t) = \int_0^t (f(s, y_n(s)) - f(s, y_{n-1}(s))) ds$$

pour $n = 1, 2, \dots$ et $0 \leq t \leq T$. On a de plus

$$y_1(t) - y_0(t) = \int_0^t f(s, a) ds.$$

Comme f est supposée continue, la fonction $s \in [0, T] \rightarrow |f(s, a)|$ est continue et donc bornée puisque l'intervalle $[0, T]$ est compact. Notons M un majorant de cette fonction, de sorte que :

$$|y_1(t) - y_0(t)| \leq TM,$$

pour tout $t \in [0, T]$. En utilisant l'hypothèse (5.2.3) et les propriétés usuelles de l'intégrale et de la norme, on majore comme suit :

$$|y_{n+1}(t) - y_n(t)| \leq L \int_0^t |y_n(s) - y_{n-1}(s)| ds.$$

Par récurrence sur $n = 1, 2, \dots$, on déduit :

$$|y_{n+1}(t) - y_n(t)| \leq MT \frac{(tL)^n}{n!}.$$

Donc, la norme du sup vérifie :

$$\|y_{n+1} - y_n\| \leq MT \frac{(TL)^n}{n!}.$$

Ainsi on a une série convergente avec

$$\sum_{n=0}^{+\infty} \|y_{n+1} - y_n\| \leq MT \exp(TL).$$

On peut donc appliquer la Remarque 5.2.10 pour en déduire que la suite $(y_n)_{n \in \mathbb{N}}$ est de Cauchy et a donc une limite y pour la convergence uniforme sur $[0, T]$. En passant à la limite dans la définition des y_n , on trouve bien la formulation intégrale de l'EDO et la démonstration du théorème 5.2.4 est ainsi terminée. ■

Remarque sur le cas linéaire 5.2.11. *Exponentielle de matrices.*

Regardons ce que donne la méthode de Picard dans le cas d'une équation linéaire autonome :

$$y' = A.y,$$

où A est une matrice carrée $m \times m$.

On est bien dans le cadre du théorème 5.2.4 puisque l'application $x \rightarrow A.x$ est Lipschitzienne. La première itération donne

$$y_1(t) = a + tA.a,$$

la seconde

$$y_2(t) = a + \int_0^t A.y_1(s)ds = a + \int_0^t A.(a + sA.a)ds = a + tA.a + \frac{t^2}{2}A^2.a,$$

d'où par une récurrence immédiate :

$$y_n(t) = a + tA.a + \frac{t^2}{2}A^2.a + \dots + \frac{t^n}{n!}A^n.a = \left(\sum_{k=0}^n \frac{t^k}{k!}A^k\right).a$$

Comme on l'a vu, la suite $(y_n)_{n \in \mathbb{N}}$ converge vers la solution y de l'EDO $y' = A.y$, qui passe par a au temps $t = 0$. Cette solution s'écrit

$$y(t) = \exp(tA).a$$

où on note

$$\exp(tA) = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^k.$$

Pour tout t réel, cette matrice carrée, appelée exponentielle de A dans le cas $t = 1$, est bien définie comme somme d'une série normalement convergente dans l'espace des matrices carrées muni de la norme matricielle $\|\cdot\|$ associée à la norme euclidienne sur \mathbb{R}^m , puisque :

$$\|\exp(tA)\| \leq \sum_{k=0}^{\infty} \frac{|t|^k}{k!} \|A\|^k = \exp(|t| \|A\|).$$

Dans le cas $m = 1$, A s'identifie à un réel et on retrouve la définition classique de l'exponentielle comme somme d'une série numérique.

Pour calculer une exponentielle de matrice, on peut donc utiliser la théorie des équations différentielles en trouvant, pour chaque $a \in \mathbb{R}^m$, la solution $\exp(tA).a$ de l'EDO linéaire autonome $y' = A.y$, avec $y(0) = a$.

Nous reviendrons sur le cas des équations linéaires à coefficients constants dans le chapitre suivant.

Le Théorème de Cauchy-Lipschitz donne aussi l'existence et l'unicité de solutions des EDO avec condition initiale dans le cas où l'intervalle de temps est infini. Rappelons que la définition de solutions dans ce cas est donnée par 5.1.2. On peut énoncer le théorème dans ce cadre-là :

Théorème 5.2.12. Soit I un intervalle de \mathbb{R}^m , $t_0 \in I$ fixé, a donné dans \mathbb{R}^m et f une fonction continue donnée de $I \times \mathbb{R}^m$ dans \mathbb{R}^m et satisfaisant 5.2.3. Alors il existe une unique fonction continue

$$t \in I \rightarrow y(t) \in \mathbb{R}^m$$

solution de l'EDO $y'(t) = f(t, y(t))$ avec la condition initiale $y(t_0) = a$.

DÉMONSTRATION : Pour tout intervalle compact J inclus dans I et contenant t_0 , le Théorème de Cauchy-Lipschitz 5.2.4 implique l'existence et l'unicité d'une solution y_J de EDO, définie sur J telle que $y_J(t_0) = a$. Grâce à la condition de Lipschitz 5.2.3, le lemme de Gronwall implique que si J' est un autre intervalle compact de I et $y_{J'}$ la solution de EDO correspondante, alors $y_J = y_{J'}$ sur $J \cap J'$. On peut donc bien définir une fonction y sur I , qui sera solution de EDO sur I au sens voulu, défini en 5.1.2. ■

5.3 Principaux résultats dans le cas localement lipschitzien

Jusqu'à maintenant, on a considéré des équations où la fonction $(t, x) \rightarrow f(t, x)$ est une fonction continue sur $[0, T] \times \mathbb{R}^m$ tout entier et est globalement Lipschitzienne en x , uniformément en t . On considère maintenant le cas plus général où f est seulement définie sur $[0, T] \times V$ où V est un ouvert connexe non vide de \mathbb{R}^m et où f est supposée continue et localement Lipschitzienne en x , uniformément en t , ce qui signifie la chose suivante :

Définition 5.3.1. Une fonction f , continue de $[0, T] \times \mathbb{R}^m$ dans \mathbb{R}^m , est dite localement lipschitzienne en x , uniformément en t si pour toute partie fermée bornée F contenue dans l'ouvert V , il existe une constante L_F telle que

$$|f(t, b) - f(t, c)| \leq L_F |b - c|, \quad \forall b, c \in F, \quad \forall t \in [0, T].$$

On notera en particulier que si une partie fermée bornée F est contenue dans l'ouvert V , elle ne touche pas le bord de V .

Lemme 5.3.2. Il suffit de vérifier cette propriété pour les boules fermées contenues dans l'ouvert V .

DÉMONSTRATION : Soit V un ouvert de \mathbb{R}^m . On suppose que f est lipschitzienne par rapport à x , uniformément en t sur toute boule fermée B incluse dans V et on note L_B la constante de Lipschitz associée à cette boule.

Soit F une partie fermée bornée de V . On rappelle que puisque F est en particulier fermée bornée dans \mathbb{R}^m , elle est compacte.

Soit $r > 0$ assez petit pour que la distance de F au bord de V soit d'au moins $3r$. Recouvrons F par l'ensemble des boules de rayon r centrées en chacun de ses points. De ce recouvrement, comme F est compacte, on peut extraire un sous-recouvrement fini par des boules fermées $B(a_i, r)$ avec $i = 1, \dots, N$ où les a_i sont dans F . Notons que les boules fermées $B(a_i, 2r)$ sont encore contenues dans V , puisque chaque a_i est dans F et donc à distance d'au moins $3r$ du bord de V .

Soient x et y deux points de F . Supposons d'abord leur distance $|x - y|$ inférieure à r . Puisque les boules $B(a_i, r)$, $i = 1, \dots, N$ recouvrent F , il existe un a_j tel que $|x - a_j| \leq r$.

Ceci implique $|y - a_i| \leq |x - a_i| + |x - y| \leq 2r$. Donc x et y sont dans la même boule fermée $B(a_i, 2r)$, elle-même contenue dans V . Notons K le sup des constantes de Lipschitz $L_{B(a_i, 2r)}$. On a donc

$$|f(t, x) - f(t, y)| \leq K|x - y|, \quad \forall t \in [0, T]$$

Si maintenant $|x - y| \geq r$, on pose $K' = 2r^{-1} \sup_{x \in F, t \in [0, T]} |f(t, x)|$. On a alors

$$|f(t, x) - f(t, y)| \leq K'r \leq K'|x - y|, \quad \forall t \in [0, T]$$

Ceci prouve bien que f est lipschitzienne par rapport à x sur F , uniformément en t , avec une constante de Lipschitz L_F inférieure ou égale au maximum de K et K' . ■

La propriété (5.3.1) est vérifiée par tous les exemples considérés dans la première section, y compris l'exemple le plus complexe de la dynamique de N corps. La raison en est donnée par la proposition suivante :

Proposition 5.3.3. *La propriété 5.3.1 est automatiquement vérifiée dès que $f(t, x)$ est continue sur $[0, T] \times V$, partout différentiable en x et que les dérivées partielles $\frac{\partial}{\partial x_i} f(t, x)$ sont continues sur $[0, T] \times V$.*

DÉMONSTRATION : Rappelons que si les dérivées partielles $\frac{\partial}{\partial x_i} f(t, x)$ sont continues sur $[0, T] \times V$, la différentielle Df de f est continue $[0, T] \times V$. Alors, puisque les boules sont convexes, pour toute boule fermée B contenue dans V et pour tous $b, c \in B$, $t \in [0, T]$, on peut écrire :

$$f(t, b) - f(t, c) = \int_0^1 Df(t, c + (b - c)s) \cdot (b - c) ds$$

On peut donc poser

$$L_B = \sup_{a \in B, t \in [0, T]} \|Df(t, a)\|,$$

où $\|\cdot\|$ est la norme d'opérateur associée à la norme euclidienne et l'inégalité précédente implique alors :

$$|f(t, b) - f(t, c)| \leq L_B |b - c|$$

f est donc lipschitzienne par rapport à x , uniformément en t sur les boules fermées de V , ce qui implique bien que cette fonction est localement lipschitzienne par rapport à x , uniformément en t sur V en appliquant le Lemme. ■

On cherche à résoudre dans ce cadre généralisé de fonctions localement lipschitzienne en x , uniformément en t , l'EDO $y'(t) = f(t, y(t))$ avec condition initiale $y(0) = y_0$. Le prix à payer pour cette généralisation est, d'abord, que y_0 doit appartenir à V . Ensuite, comme le montre l'exemple $m = 1$, $f(t, x) = x^2$, on ne peut pas s'attendre à pouvoir résoudre l'EDO sur l'intervalle de temps $[0, T]$ tout entier. On utilise donc la notion de solution locale définie sur un sous intervalle de temps $[0, \tau]$, par opposition à solution globale si elle est définie sur tout l'intervalle de temps $[0, T]$:

Définition 5.3.4. On dit qu'une fonction continue $t \rightarrow y(t)$ est solution locale de l'EDO $y'(t) = f(t, y(t))$ avec condition initiale $y(0) = y_0$ si elle est définie sur un intervalle $[0, \tau]$, avec $0 < \tau \leq T$, à valeurs dans V et qu'elle vérifie l'(EDO), par exemple sous la forme intégrale

$$y(t) = y_0 + \int_0^t f(s, y(s)) ds, \quad \forall t \in [0, \tau].$$

La première propriété qu'assure l'hypothèse faite sur f d'être localement lipschitzienne en x , uniformément en t , est le résultat suivant d'unicité et de stabilité par rapport aux données initiales :

Proposition 5.3.5. Soient y et z deux solutions locales d'une EDO $y'(t) = f(t, y(t))$ où f est localement lipschitzienne en x , uniformément en t , définies respectivement sur les intervalles $[0, \tau]$ et $[0, \sigma]$, avec les conditions initiales $y(0) = y_0$ et $z(0) = z_0$ dans V . On suppose $\tau \leq \sigma$. Alors on a l'estimation

$$|y(t) - z(t)| \leq |y_0 - z_0| \exp(L_K t), \quad \forall t \in [0, \tau],$$

où K est la réunion de $y([0, \tau])$ et $z([0, \tau])$. En particulier, si $y_0 = z_0$ alors y et z coïncident sur $[0, \tau]$.

DÉMONSTRATION : Notons d'abord que comme y et z sont continues de $[0, \tau]$ à valeurs dans V , leurs images sont des compacts inclus dans V et il en est donc de même pour leur réunion K . Soit L_K la constante de Lipschitz de f associée à K . On procède exactement comme dans le cas globalement lipschitzien :

$$\begin{aligned} \frac{d}{dt} |y(t) - z(t)|^2 &= 2 \langle y(t) - z(t), f(t, y(t)) - f(t, z(t)) \rangle \\ &\leq 2L_K |y(t) - z(t)|^2, \end{aligned}$$

Par le lemme de Gronwall 5.2.7, ceci implique alors :

$$|y(t) - z(t)| \leq |y_0 - z_0| \exp(L_K t)$$

ce qui prouve la Proposition. ■

Ce premier résultat d'unicité conduit naturellement au concept de solution maximale pour une donnée initiale y_0 fixée.

Définition 5.3.6. Si l'équation différentielle EDO a au moins une solution locale avec la condition initiale $y(0) = y_0$, on définit le temps maximal d'existence T^* comme la borne supérieure des $\tau \in]0, T]$ pour lesquels on peut définir une solution locale y sur $[0, \tau]$ partant de y_0 au temps 0. On peut alors définir une unique solution maximale partant de y_0 sur l'intervalle $[0, T^*[$.

Bien sur, T^* peut être strictement plus petit que T et dépendre alors de y_0 , comme le montre bien l'exemple élémentaire $m = 1$, $V = \mathbb{R}$, $f(t, x) = x^2$ où on a $T^* = \min(T, y_0^{-1})$ si $y_0 > 0$.

On peut montrer que si $T^* < T$ alors, nécessairement, ou $|y(t)|$ tend vers l'infini ou la distance de $y(t)$ au bord de l'ouvert V tend vers 0 lorsque t s'approche de T^* .

Théorème 5.3.7. (Théorème de Cauchy-Lipschitz local) Soit f une fonction continue de $[0, T] \times V$ dans \mathbb{R}^m et localement Lipschitzienne en x , uniformément en t . L'EDO $y'(t) = f(t, y(t))$ avec condition initiale $y(0) = y_0$ admet au moins une solution locale.

DÉMONSTRATION : Pour prouver l'existence de cette solution locale, il suffit d'adapter la preuve du cas globalement lipschitzien utilisant la méthode des approximations successives. Prenons $r > 0$ tel que la boule fermée B centrée en y_0 de rayon r soit incluse dans V .

Soient d'autre part $M = \sup_{s \in [0, T], x \in B} |f(s, x)|$ et $\tau > 0$ tel que $\tau M \leq r$.

Faisons l'hypothèse de récurrence que l'approximation $y_n(t)$ de la méthode de Picard à l'étape n soit définie dans l'intervalle $t \in [0, \tau]$, et prenne ses valeurs dans B . A l'étape suivante, on a

$$y_{n+1}(t) - y_0 = \int_0^t f(s, y_n(s)) ds$$

et donc

$$|y_{n+1}(t) - y_0| \leq \tau M,$$

L'hypothèse de récurrence est bien vérifiée à l'étape $n + 1$. Ainsi, le procédé est bien défini dans la boule B et on peut alors reprendre exactement le raisonnement du cas globalement lipschitzien dans l'espace $\mathcal{C}([0, \tau]; B)$, qui est complet, avec pour constante de Lipschitz L_B . On obtient notamment, pour $n = 1, 2, \dots$,

$$|y_{n+1}(t) - y_n(t)| \leq L_B \int_0^t |y_n(s) - y_{n-1}(s)| ds.$$

■

Dans de nombreuses applications, où f est seulement localement Lipschitzienne, on parvient tout de même à montrer l'existence de solutions globales grâce à l'existence d'une fonction de Liapounov associée à l'EDO :

Définition 5.3.8. Soit une fonction U de $V \subset \mathbb{R}^m$ dans $[0, +\infty[$, continument différentiable. On dit que U est une fonction de Liapounov pour l'EDO $y'(t) = f(t, y(t))$ si :

- 1) U est une fonction strictement croissante de $|x|$
- 2) pour tout $M \in \mathbb{R}$, l'ensemble

$$\{x \in V, U(x) \leq M\}$$

est un compact inclus dans V

- 3) Il existe deux constantes $\alpha \geq 0$ et $\beta \geq 0$ telles que

$$DU(x).f(t, x) \leq \alpha + \beta U(x), \quad \forall t \in [0, T], \quad \forall x \in V.$$

où $DU(x)$ représente la différentielle de la fonction U au point x .

Définition 5.3.9. Lorsque dans la condition 3), $\alpha = \beta = 0$, on dira que U est une fonction de Liapounov au sens strict pour l'EDO.

Exemples. Dans beaucoup de cas, il suffit de prendre $U(x) = |x|^2$, auquel cas on recherche une inégalité de la forme :

$$2f(t, x).x \leq \alpha + \beta|x|^2, \quad \forall t \in [0, T], \quad \forall x \in V.$$

C'est le cas de l'exemple $m = 1$, $V = \mathbb{R}$, $f(t, x) = -x^3$, qui ne rentre pas dans le cadre pas globalement lipschitzien.

Notons que la condition 3) est automatiquement satisfaite avec la fonction $U(x) = |x|^2$, s'il existe une constante K telle que

$$|f(t, x)| \leq (1 + |x|)K \quad \forall t \in [0, T], \quad \forall x \in V$$

et cette dernière condition est elle-même automatiquement satisfaite si $f(t, x)$ est globalement lipschitzienne en x , uniformément en t .

On retrouve ainsi le cas du Théorème de Cauchy-Lipschitz, 5.2.4 .

On verra cependant plus loin deux classes fondamentales d'EDO pour lesquelles des fonctions de Liapounov autres que $U(x) = |x|^2$ apparaissent naturellement.

La principale conséquence de l'existence d'une fonction de Liapounov est le résultat suivant :

Proposition 5.3.10. *S'il existe une fonction de Liapounov généralisée sur V pour l'EDO $y'(t) = f(t, y(t))$, où f est localement lipschitzienne en x uniformément en t , alors pour toute donnée initiale $y_0 \in V$, il existe une solution globale sur l'intervalle $[0, T]$ tout entier.*

DÉMONSTRATION : L'idée est de se ramener au cas globalement lipschitzien par une opération de coupure. Etant donné $M > 0$ fixé, on introduit

$$K = \{x \in V, U(x) \leq M\}, \quad K' = \{x \in V, U(x) \leq 2M\}$$

et la fonction de coupure

$$\theta(x) = \frac{d(x, \mathbb{R}^m \setminus K')}{d(x, \mathbb{R}^m \setminus K') + d(x, K)}$$

où on note par $d(x, A)$ la distance d'un point $x \in \mathbb{R}^m$ à une partie non vide A de \mathbb{R}^m .

Cette fonction θ a la propriété d'être définie et globalement lipschitzienne sur \mathbb{R}^m , à valeurs dans $[0, 1]$, de valoir 0 sur $\mathbb{R}^m \setminus K'$ et 1 sur K . On introduit alors le champ de vecteur tronqué

$$f_\theta(t, x) = \theta(x)f(t, x)$$

qu'on étend à $[0, T] \times \mathbb{R}^m$ par la valeur 0 si x est hors de V . Cela fait de $f_\theta(t, x)$ un champ globalement lipschitzien en $x \in \mathbb{R}^m$, uniformément en $t \in [0, T]$.

En effet, fixons b, c dans \mathbb{R}^m , t dans $[0, T]$ et posons

$$r = |f_\theta(t, b) - f_\theta(t, c)|$$

Si b et c sont hors de K' , on trouve $r = 0$ par définition (car $\theta(b) = \theta(c) = 0$). Si b est dans K' et c hors de K' , on a $\theta(c) = 0$ et donc :

$$\begin{aligned} r &= |f(t, b)\theta(b) - f(t, c)\theta(c)| = |f(t, b)\theta(b)| \\ &\leq C|\theta(b)| = C|\theta(b) - \theta(c)| \leq CL_\theta|b - c|, \end{aligned}$$

où on note

$$C = \sup_{s \in [0, T], x \in K'} |f(s, x)|$$

et L_θ la constante de Lipschitz de θ . Enfin, si b et c sont dans K' , on a

$$r \leq |(f(t, b) - f(t, c))\theta(b)| + |f(t, c)(\theta(b) - \theta(c))| \leq (L_{K'} + CL_\theta)|b - c|,$$

où $L_{K'}$ est la constante de Lipschitz de f en $x \in K'$. Ainsi, f_θ a une constante de Lipschitz en x uniformément bornée par $L_{K'} + CL_\theta$.

Considérons maintenant l'unique solution globale

$$t \in [0, T] \rightarrow y(t) \in \mathbb{R}^m$$

de l'EDO $y'(t) = f_\theta(t, y(t))$ ayant y_0 pour donnée initiale. Cette solution est bien globale, puisque f_θ est globalement lipschitzienne en $x \in \mathbb{R}^m$, uniformément en $t \in [0, T]$.

Montrons que si on choisit $M \geq U(y_0)$ assez grand, on a

$$U(y(t)) \leq M \quad \forall t \in [0, T].$$

Supposons qu'il existe $0 < \tau$ tel que $\tau = \inf\{t \in [0, T] \mid U(y(t)) \geq M\}$. Alors, on effectue le calcul suivant :

$$\begin{aligned} \frac{d}{dt}U(y(t)) &= DU(y(t)) \cdot f_\theta(t, y(t)) \\ &= DU(y(t)) \cdot f(t, y(t))\theta(y(t)) \end{aligned}$$

par définition de f_θ

$$\leq (\alpha + \beta U(y(t)))\theta(y(t))$$

par hypothèse et parce que $\theta \geq 0$

$$\leq \alpha + \beta U(y(t))$$

parce que $\theta \leq 1$, α, β et U sont positives.

Par le lemme de Gronwall, on en déduit l'inégalité

$$U(y(t)) \leq U(y_0) + \alpha t \quad \forall t \in [0, \tau] \text{ si } \beta = 0$$

$$U(y(t)) \leq U(y_0) \exp(\beta t) + \alpha \frac{\exp(\beta t) - 1}{\beta}, \quad \forall t \in [0, \tau] \text{ si } \beta > 0$$

Dans les deux cas, si on a pris le soin de choisir M assez grand en fonction de $U(y_0)$, α , β et T , à savoir

$$M > U(y_0) + \alpha T \text{ si } \beta = 0$$

$$M > U(y_0) \exp(\beta T) + \alpha \frac{\exp(\beta T) - 1}{\beta} \text{ si } \beta > 0$$

on aura alors obtenu

$$U(y(\tau)) < M$$

ce qui est contradictoire avec la définition de τ . Ainsi on a prouvé

$$U(y(t)) \leq M \quad \forall t \in [0, T].$$

ce qui assure $y(t) \in K$, pour tout $t \in [0, T]$, par définition de K .

Comme la fonction de coupure θ vaut 1 sur K , ceci prouve qu'en fait $y'(t) = f(t, y(t))$ pour tout t , ce qui veut dire que y est en fait solution de l'EDO $y'(t) = f(t, y(t))$ avec y_0 pour donnée initiale. Elle en est donc l'unique solution globale et la preuve est achevée. ■

Exemple 5.3.11. On considère le système différentiel, défini pour $t \geq 0$ par :

$$(P) \quad \begin{cases} u'(t) + u^3(t) = v(t) \\ v'(t) + v^3(t) = u(t) \\ u(0) = u_0, v(0) = v_0 \end{cases}$$

La fonction $f : (x, y) \rightarrow (-x^3 + y, -y^3 + x)$ est autonome, de classe C^1 de \mathbb{R}^2 dans lui-même. Elle est donc localement lipschitzienne. Pour toutes conditions initiales (u_0, v_0) , il existe donc une solution locale.

La fonction $U(x, y) = x^2 + y^2$ est une fonction de Liapounov pour (P) . En effet, les lignes de niveau sont des disques fermés de \mathbb{R}^2 et donc sont compacts. De plus, on a :

$$U'(x, y)f(x, y) = 2x(-x^3 + y) + 2y(-y^3 + x) = -2x^4 - 2y^4 + 4xy \leq 2(x^2 + y^2) = 2U(x, y)$$

Il existe donc des solutions globales sur tout compact donc sur $[0, +\infty)$ pour toutes conditions initiales.

Exemples 5.3.12. Equations de type gradient.

Définition 5.3.13. Soit U une fonction de $V \subset \mathbb{R}^m \rightarrow \mathbb{R}$ et $x \in V$. On appelle gradient de U au point x et on note $\nabla U(x)$, le vecteur de \mathbb{R}^m ayant pour coordonnées les dérivées partielles de U au point x , c'est à dire :

$$\nabla U(x) = \begin{pmatrix} \frac{\partial U}{\partial x_1}(x) \\ \vdots \\ \frac{\partial U}{\partial x_m}(x) \end{pmatrix}$$

Remarquons que la différentielle $DU(x)$ de U au point x est une application linéaire de \mathbb{R}^m dans \mathbb{R} et qu'elle est représentée par la matrice ligne :

$$DU(x) = \left(\frac{\partial U}{\partial x_1}(x) \quad \dots \quad \frac{\partial U}{\partial x_m}(x) \right)$$

Définition 5.3.14. Une EDO autonome est de type gradient s'il existe une fonction, deux fois continument différentiable, U de V dans \mathbb{R} telle que

$$f(x) = -\nabla U(x) \quad \forall x \in V.$$

Tous les systèmes autonomes, avec f continument différentiable, ne sont pas nécessairement de type gradient dès que $m > 1$. Les systèmes hamiltoniens, que l'on va bientôt considérer, en sont des contre-exemples typiques. Le contre-exemple le plus simple est donnée par équation linéarisée du pendule : $m = 2$, $V = \mathbb{R}^2$, $f(x) = (x_2, -\frac{g}{L}x_1)$.

Si U est une fonction croissante de $|x|$ et a la propriété de compacité de ses ensembles de niveau 5.3.8 1), elle constitue une fonction de Liapounov au sens strict, puisque

$$DU(x).f(x) = -|\nabla U(x)|^2 \leq 0 \quad \forall x \in V.$$

Ainsi, pour toute solution de l'EDO, on a l'estimation

$$\frac{d}{dt}(U(y(t))) = -|\nabla U(y(t))|^2 \leq 0 \quad \forall t \in [0, T].$$

Comme exemple, on citera le cas des systèmes linéaires autonomes

$$y' = A.y$$

où A est une matrice symétrique définie négative, ce qui correspond à poser $f(x) = -\nabla U(x)$ avec $U(x) = -\frac{1}{2}\langle Ax, x \rangle$.

Exemple 5.3.15. *Equations de type hamiltonien.*

Lorsque la dimension $m = 2m'$ est paire, on rencontre souvent la structure suivante, dite hamiltonienne

Définition 5.3.16. *On note la variable d'espace $x = (q, p)$, q désignant les m premières coordonnées (dites de position en mécanique) et p les m' suivantes (dites d'impulsion). On pose :*

$$f(x) = J\nabla U(x) \quad \forall x \in V,$$

où J est l'opérateur linéaire de \mathbb{R}^m dans \mathbb{R}^m

$$J(q, p) = (-p, q).$$

et où U est une fonction, deux fois continument différentiable, de $V \subset \mathbb{R}^m$ dans \mathbb{R} . On dira que U est l'hamiltonien de l'EDO : $y'(t) = f(y(t))$.

Là encore si U est une fonction croissante de $|x|$ et vérifie la propriété de compacité de ses ensembles de niveau 5.3.8 1), elle constitue une fonction de Liapounov pour l'EDO : $y'(t) = f(y(t))$ puisque

$$DU(x).f(x) = DU(x).J\nabla U(x) = \langle \nabla U(x), J\nabla U(x) \rangle = 0,$$

par antisymétrie de l'opérateur J .

En fait, on a la propriété de conservation de l'hamiltonien car pour toute solution $y(t)$ de l'EDO on aura $U(y(t)) = U(y_0)$, autrement dit l'hamiltonien est une constante d'intégration.

Comme exemple, prenons le cas de la dynamique d'une particule de masse unité dans un potentiel φ . On a alors $m' = 3$, $m = 6$ et l'hamiltonien est alors donné par

$$U(x) = \frac{1}{2}|p|^2 + \varphi(q).$$

La croissance en fonction de $|x|$ et la propriété de compacité des ensembles de niveau 5.3.8 1) seront satisfaite ou non en fonction des propriétés du potentiel φ . Ce sera le cas par exemple du potentiel quadratique $\varphi(q) = \frac{1}{2}|q|^2$ (oscillateur harmonique).

5.4 Appendice sur le schéma d'Euler

Cet appendice a pour but de montrer directement que les approximations fournies par le schéma d'Euler forment une suite de Cauchy sans supposer l'existence d'une solution, ce qui permet du même coup de montrer qu'il existe une solution dans le théorème de Cauchy-Lipschitz 5.2.4 *sans* utiliser la méthode des approximations successives de Picard.

Pour alléger les notations, nous nous limiterons à la preuve dans le cas autonome où la donnée f ne dépend pas de la variable temps t .

Nous allons montrer que la suite des approximations d'Euler $y_h(t)$ obtenues pour les h de la forme $h = h_0 2^{-k}$ pour $k = 1, 2, \dots$, et $h_0 > 0$ fixé, est une suite de Cauchy dans l'espace $\mathcal{C}([0, T], E)$. Cela nous impliquera bien que la suite converge uniformément sur $I = [0, T]$. Montrons d'abord que l'on peut borner $y_h(t)$ uniformément en h et t . A cause de la condition de Lipschitz, il existe une constante K telle que

$$|f(c)| \leq K + L|c|, \quad \forall c \in E.$$

Il suffit de poser $b = 0$ dans 5.2.3, d'utiliser l'inégalité du triangle et de prendre $K = |f(0)|$. A partir du schéma d'Euler, on obtient

$$|y_h(nh + h)| \leq |y_h(nh)| + h|f(y_h(nh))|$$

et donc, par l'inégalité ci-dessus,

$$|y_h(nh + h)| \leq hK + (1 + hL)|y_h(nh)|.$$

Cette inégalité de récurrence donne, selon la Proposition 5.2.8,

$$|y_h(nh)| + K/L \leq (|a| + K/L)(1 + hL)^n \leq (|a| + K/L) \exp(nhL).$$

Ainsi, pour tout $T > 0$ fixé, on a

$$\sup_{nh \leq T} |y_h(nh)| \leq (|a| + K/L) \exp(TL) - K/L.$$

En réutilisant l'inégalité ci-dessus, on obtient aussi

$$\sup_{nh \leq T} |f(y_h(nh))| \leq (L|a| + K) \exp(TL).$$

L'idée est maintenant de comparer les deux solutions approchées de le schéma d'Euler pour les pas de temps respectifs h et $2h$. Considérons le temps nh où n est un entier positif ou nul fixé tel que $nh \leq T$. Sur l'intervalle $[nh, nh + 2h]$, y_{2h} est affine et

$$y_{2h}(nh + 2h) = y_{2h}(nh) + 2hf(y_{2h}(nh)).$$

On a donc

$$y_{2h}(nh + h) = y_{2h}(nh) + hf(y_{2h}(nh))$$

et

$$y_{2h}(nh + 2h) = y_{2h}(nh + h) + hf(y_{2h}(nh + h))$$

Sur ce même intervalle y_h est affine sur chaque demi-intervalle $[nh, nh + h]$ et $[nh + h, nh + 2h]$, avec

$$y_h(nh + h) = y_h(nh) + hf(y_h(nh)),$$

$$y_h(nh + 2h) = y_h(nh + h) + hf(y_h(nh + h)).$$

Parce que les deux fonctions sont affines sur les deux demi-intervalles, on voit, à l'aide de l'inégalité du triangle, que le maximum de la norme de leur différence sur l'intervalle $[nh, nh + 2h]$ est donné par le maximum des trois nombres suivants

$$e_0 = |y_h(nh) - y_{2h}(nh)|, \quad e_1 = |y_h(nh + h) - y_{2h}(nh + h)|, \quad e_2 = |y_h(nh + 2h) - y_{2h}(nh + 2h)|.$$

En utilisant la condition de Lipschitz sur f et les formules précédentes on peut majorer

$$e_1 \leq |y_h(nh) - y_{2h}(nh)|(1 + hL).$$

On a donc

$$e_1 \leq (1 + hL)e_0.$$

Pour majorer e_2 , on observe que

$$y_h(nh + 2h) - y_{2h}(nh + 2h) = y_{2h}(nh + h) - y_h(nh + h) + h(f(y_h(nh + h)) - f(y_{2h}(nh))).$$

Donc

$$\begin{aligned} e_2 &\leq e_1 + hL|y_h(nh + h) - y_{2h}(nh)| \\ &\leq e_1 + Lh|y_h(nh + h) - y_h(nh)| + Lh|y_h(nh) - y_{2h}(nh)| \end{aligned}$$

par l'inégalité du triangle

$$\begin{aligned} &\leq e_1 + Lh|hf(y_h(nh))| + Lhe_0 \\ &\leq e_1 + Lh^2(L|a| + K) \exp(TL) + Lhe_0 \end{aligned}$$

en utilisant la relation $|f(c)| \leq K + L|c|$ et donc

$$e_2 \leq (1 + 2hL)e_0 + Lh^2(L|a| + K) \exp(TL)$$

en utilisant $e_1 \leq (1 + hL)e_0$.

On a donc obtenu , en notant

$$e_h(t) = |y_h(t) - y_{2h}(t)|$$

que

$$\sup_{nh \leq t \leq nh+2h} e_h(t) \leq (1 + 2hL)e_h(nh) + Lh^2(L|a| + K) \exp(TL)$$

et, en particulier, pour $t = nh + h$:

$$e_h(nh + 2h) \leq (1 + 2hL)e_h(nh) + Lh^2(L|a| + K) \exp(TL)$$

En utilisant de nouveau la Proposition 5.2.8, et compte tenu de ce que $e_h(0) = 0$ (puisque'on a $y_{2h}(0) = y_h(0) = a$), on aboutit finalement à

$$\sup_{nh \leq T} e_h(nh) \leq hQ(L, K, \exp(TL))/2$$

où Q est donné par

$$Q(L, K, \exp(TL)) = (L|a| + K) \exp(TL)(\exp(TL) - 1).$$

Ainsi

$$\sup_{0 \leq t \leq T} |y_h(t) - y_{2h}(t)| \leq hQ(L, K, \exp(TL))/2.$$

Si l'on prend la suite des pas de temps $h_k = h_0 2^{-k}$ pour $k = 0, 1, 2, \dots$, la suite des y_{h_k} est une suite de Cauchy pour la norme du sup. En effet, la série

$$\sum_k \|y_{h_k} - y_{h_{k+1}}\| \leq h_0/2Q(L, K, \exp(TL)) \sum_k 2^{-k}$$

est convergente et on peut appliquer le critère des séries normalement convergentes.

Ainsi il existe une fonction continue $t \in [0, T] \rightarrow y(t)$ vers laquelle la suite $(y_{h_k})_{k \in \mathbb{N}}$ converge uniformément. On a de plus

$$\sup_{0 \leq t \leq T} |y_{h_k}(t) - y(t)| \leq h_0 2^{-k} Q(L, K, \exp(TL)),$$

pour tout $k = 0, 1, \dots$, en utilisant une majoration du reste d'une série, ce qui nous donne en particulier, pour $k = 0$, l'estimation d'erreur

$$\sup_{0 \leq t \leq T} |y_h(t) - y(t)| \leq hQ(L, K, \exp(TL)).$$

Il reste ensuite à prouver que y est bien l'unique solution de l'EDO avec a pour donnée initiale. On aura ainsi prouvé directement l'existence d'une solution et la convergence du schéma d'Euler.

Or, on a, en utilisant la deuxième formulation du schéma d'Euler,

$$y_h(t) = a + h \sum_{kh < t_h} f(y_h(kh)) + (t - t_h)f(y_h(t_h)).$$

En passant à la limite, on retrouve la formulation intégrale de l'équation différentielle, ce qui termine la preuve. ■

Remarque. Pour ce dernier passage à la limite, on pourra utiliser comme étape intermédiaire pour chaque $k = 0, 1, \dots$ tel que $kh < t$, les majorations :

$$\begin{aligned} & \left| \int_{kh}^{(k+1)h} f(y_h(s)) ds - hf(y_h(kh)) \right| \\ & \leq \int_{kh}^{(k+1)h} |f(y_h(s)) - f(y_h(kh))| ds \\ & \leq L \int_{kh}^{(k+1)h} |y_h(s) - y_h(kh)| ds \\ & \leq Lh^2 |f(y_h(kh))| \leq Lh^2 (L|a| + K) \exp(TL) \end{aligned}$$

5.5 Autres énoncés

Il y a de multiples énoncés possibles d'existence de solutions d'une équation différentielle d'ordre 1. Retenons en deux, l'un concernant l'existence de solutions globales dans un cadre plus général que le cas Lipschitzien, et l'autre concernant l'existence de solutions locales dans un cadre utilisable en géométrie différentielle. On admettra ces résultats sans preuves.

Theoreme 5.5.1. Soit I un intervalle de \mathbb{R} , t_0 et y_0 donnés respectivement dans I et dans $E = \mathbb{R}^m$ et soit f continue de $I \times E$ dans E . On suppose qu'il existe une constante C telle que

$$|f(t, b)| \leq (1 + |b|)C,$$

pour tous $t \in I$ et $b \in E$, et que, pour chaque $R > 0$, il existe une constante L_R telle que

$$|f(t, b) - f(t, c)| \leq L_R |b - c|,$$

pour tous $t \in I$, $(b, c) \in E$ tels que $|b| \leq R$, $|c| \leq R$. Alors, il existe une unique solution sur I de l'EDO $y' = f(t, y)$ passant par y_0 à l'instant t_0 .

Remarque. Lorsque f vérifie les hypothèses du théorème précédent, on dit que f est localement Lipschitzienne et a une croissance linéaire à l'infini.

Il est assez facile d'adapter à ce cas la démonstration du théorème 5.2.4 si on l'étudie attentivement.

Theoreme 5.5.2. Soit $I =]a, b[$ un intervalle ouvert de \mathbb{R} , V un ouvert de $E = \mathbb{R}^m$. Soient t_0 et y_0 donnés respectivement dans I et dans V et soit f continue de $I \times E$ dans E . On suppose que $f(t, x)$ est dérivable par rapport à x et que les dérivées partielles $\frac{\partial f(t, x)}{\partial x_i}$ sont continues. Alors il existe un intervalle $J =]c, d[$ contenant t_0 et inclus dans I et une solution y de l'EDO $y' = f(t, y)$ passant par y_0 au temps t_0 , qui est l'unique solution sur l'intervalle J .

Remarque. Il y a clairement non-unicité de l'intervalle J , tout intervalle ouvert plus petit contenant t_0 convient aussi bien. Néanmoins, on peut définir le plus grand d'entre eux, en prenant toutes leurs réunions, qu'on appelle intervalle maximal. La solution y correspondante est dite maximale.

On peut alors montrer que l'on a forcément l'une des propriétés suivantes à droite de l'intervalle J : ou bien $d = b$, ou bien $d < b$ et $|y(t)| \rightarrow +\infty$ quand $t \rightarrow d$, ou bien, si V n'est pas égal à E tout entier, $d < b$ et $y(t)$ sort de l'ouvert V lorsque $t \rightarrow d$, c'est-à-dire que sa distance au complémentaire de V dans E tend vers zéro. On a la propriété symétrique du côté gauche de l'intervalle J : ou bien $c = a$, ou bien $c > a$ et $|y(t)| \rightarrow +\infty$ etc....

Si l'intervalle de temps est trop long, on se ramène au cas précédent en segmentant l'intervalle en sous-intervalles et en résolvant l'EDO de proche en proche. En effet :

Proposition 5.5.3. (Recollement) Soit y et f à valeurs dans \mathbb{R}^m définies respectivement sur $[0, T]$ et $[0, T] \times V$. Soit $K > 0$ entier et $I_k = [(k-1)T/K, kT/K]$ pour $k = 1, \dots, K$. Si

y est continue sur $[0, T]$ et est solution de l'EDO $y' = f(t, y)$ sur chaque sous-intervalle I_k , alors y est solution de la même EDO sur l'intervalle $[0, T]$.

DÉMONSTRATION : La preuve est une conséquence directe de la formulation intégrale de l'EDO. Détaillons dans le cas le plus simple $K = 2$. Pour $0 \leq t \leq T/2$, on a directement

$$y(t) = a + \int_0^t f(y(s))ds,$$

puisque y est solution de l'EDO sur $[0, T/2]$ par hypothèse. En particulier

$$y(T/2) = a + \int_0^{T/2} f(y(s))ds.$$

Pour $T/2 \leq t \leq T$, on a

$$y(t) = y(T/2) + \int_{T/2}^t f(y(s))ds,$$

puisque y est solution de l'EDO sur $[T/2, T]$ par hypothèse et passe par $y(T/2)$ au temps $T/2$. Par substitution, on obtient donc bien la formulation intégrale de la solution pour $T/2 \leq t \leq T$. Le cas $K > 2$ se démontre facilement en généralisant cet argument. ■

Remarque. Il existe des théorèmes d'existence de solutions d'équations différentielles sans unicité. On pourra par exemple trouver deux solutions distinctes de l'EDO $y' = \sqrt[3]{y}$ avec la condition initiale $y(0) = 0$.

6

EQUATIONS DIFFÉRENTIELLES LINÉAIRES

Dans ce chapitre, on note $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ l'espace vectoriel des opérateurs linéaires de \mathbb{R}^m dans lui-même. On identifie un opérateur linéaire de \mathbb{R}^m dans lui-même avec la matrice qui le représente dans la base canonique de \mathbb{R}^m .

Comme dans les chapitres précédents, \mathbb{R}^m est muni de la norme euclidienne. On rappelle que le norme d'un élément $A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ est alors donnée par :

$$\|A\| = \sup_{|x| \leq 1} |Ax|$$

6.1 Equations différentielles linéaires sans deuxième membre

Définition 6.1.1. Soit I un intervalle quelconque de \mathbb{R} . On appelle *équation différentielle linéaire homogène* une équation différentielle de la forme :

$$X'(t) = A(t)X(t) \text{ avec } \forall t \in I, A(t) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m).$$

Théorème 6.1.2. (Solution d'une équation différentielle linéaire homogène)

Soit $x_0 \in \mathbb{R}^m$. Si l'application $t \rightarrow A(t)$ est continue de I dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$, l'équation différentielle linéaire

$$EDO \quad X'(t) = A(t)X(t)$$

admet une et une seule solution X , définie sur I , à valeurs dans \mathbb{R}^m et vérifiant la condition initiale : $X(t_0) = x_0$.

DÉMONSTRATION : On pose $f(t, x) = A(t)x$

Soit J un sous intervalle compact de I tel que $t_0 \in I$. En posant

$$k = \sup\{\|A(t)\| \mid t \in J\}$$

où $\|\cdot\|$ est la norme d'opérateur dans $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$, on obtient :

$$\forall t \in J, |f(t, x_1) - f(t, x_2)| \leq k |x_1 - x_2|$$

On peut donc appliquer le Théorème d'existence et d'unicité des solutions d'équations différentielles de Cauchy-Lipschitz, 5.2.4 si I est compact et son extension aux intervalles quelconques, 5.2.12 si I n'est pas compact. ■

Remarque 6.1.3. L'application qui, à une solution X d'une équation différentielle linéaire EDO sans deuxième membre, définie sur un intervalle I , associe $X(t_0)$ est linéaire et bijective.

Cette remarque permet d'énoncer le corollaire suivant :

Corollaire 6.1.4. 1) L'ensemble \mathcal{S} des solutions d'une équation différentielle linéaire à valeurs dans \mathbb{R}^m est un espace vectoriel algébriquement isomorphe à \mathbb{R}^m .

2) Si I est un intervalle fermé borné, l'application qui à x_0 associe la solution X telle que $X(t_0) = x_0$ vérifie :

$$|X(t)| \leq |x_0| e^{k|t-t_0|} \text{ avec } k = \sup_{t \in I} \|A(t)\|$$

et par suite cette application est continue.

DÉMONSTRATION : La majoration est une conséquence immédiate du lemme de Gronwall 5.2.7 appliqué à la fonction $t \rightarrow |X(t)|^2$. ■

Définition 6.1.5. (Résolvante)

On appelle résolvante de l'équation différentielle linéaire EDO $X'(t) = A(t)X(t)$, définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R}^m et telle que $t \rightarrow A(t)$ soit continue, l'application $R : I \times I \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ telle que

$$\forall t_1, t_2 \in I, X(t_2) = R(t_2, t_1)X(t_1)$$

On peut énoncer les propriétés fondamentales de la résolvante :

Théorème 6.1.6. (Propriétés de la résolvante)

- i) $R(t_3, t_2) \circ R(t_2, t_1) = R(t_3, t_1)$
- ii) $R(t_1, t_1) = Id$, où I_E est l'identité de E
- iii) $R(t_2, t_1) = (R(t_1, t_2))^{-1}$.

DÉMONSTRATION : i) Pour toute solution X de l'équation différentielle EDO, on a :

$$X(t_3) = R(t_3, t_2)X(t_2) = R(t_3, t_2)R(t_2, t_1)X(t_1)$$

et aussi

$$X(t_3) = R(t_3, t_1)X(t_1)$$

d'où la première relation puisque $X(t_1)$ est arbitraire.

Les relations ii) et iii) découlent directement de la définition de la résolvante R . ■

On remarquera en particulier que si on a la condition initiale $X(t_0) = x_0$, alors

$$X(t) = R(t, t_0)x_0$$

Théorème 6.1.7. (*Caractérisation de la résolvante*)

On considère l'équation différentielle linéaire EDO $X'(t) = A(t)X(t)$, définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R}^m et telle que $t \rightarrow A(t)$ soit continue et on fixe $\theta \in I$.

La résolvante R de EDO est l'unique solution de l'équation différentielle dans $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$:

$$R'_t(t, \theta) = A(t) \circ R(t, \theta), \text{ avec } R(\theta, \theta) = I_{\mathbb{R}^m}$$

DÉMONSTRATION : Considérons à priori l'équation différentielle :

$$R'_t(t, \theta) = A(t) \circ R(t, \theta)$$

C'est une équation différentielle linéaire à valeurs dans $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$. En effet, l'opérateur

$$\Phi(t) : U \rightarrow A(t) \circ U$$

est un endomorphisme de $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$.

De plus, par un calcul facile, on a :

$$\|\Phi(t)\| = \|A(t)\| \text{ et aussi } \|\Phi(t_1) - \Phi(t_2)\| = \|A(t_1) - A(t_2)\|$$

Par suite, si I est fini, le théorème de Cauchy-Lipschitz implique que l'équation différentielle considérée admet une solution unique sur I vérifiant $R(\theta, \theta) = I_{\mathbb{R}^m}$.

Pour $X(\theta) \in \mathbb{R}^m$, posons $X(t) = R(t, \theta)X(\theta)$, alors :

$$X'(t) = R'_t(t, \theta)X(\theta) = A(t) \circ R(t, \theta)X(\theta) = A(t)X(t)$$

La solution $R(t, \theta)$ de l'équation différentielle considérée au départ est donc bien la valeur en (t, θ) de la résolvante de l'équation

$$X'(t) = A(t)X(t)$$

Si I n'est pas fini, on conclut comme pour le théorème 6.1.2 ■

La démonstration précédente est valable pour des équations différentielles linéaires à valeur dans un espace de Banach E , en particulier de dimension infinie. Dans le cas de la dimension finie, on a une démonstration plus simple, présentée ci-dessous. On utilise le fait que, en dimension finie, si $t \rightarrow R(t, t_0)x_0$ est dérivable pour tout $x_0 \in E$, alors $t \rightarrow R(t, t_0)$ est dérivable (c'est faux en dimension quelconque).

Remarque 6.1.8. Si l'on sait que $t \rightarrow R(t, \theta)$ est une fonction dérivable par rapport à $t \in I$ (ce qui est réalisé en particulier dans le cas où l'équation considérée est à valeurs dans \mathbb{R}^m), on peut obtenir le résultat précédent plus simplement.

DÉMONSTRATION : On dérive par rapport à t la relation

$$X(t) = R(t, \theta)X(\theta)$$

et on obtient :

$$X'(t) = R'_t(t, \theta)X(\theta)$$

Comme par ailleurs,

$$X'(t) = A(t)X(t) = A(t) \circ R(t, \theta)X(\theta)$$

et que $X(\theta)$ est arbitraire, on obtient bien

$$R'_t(t, \theta) = A(t) \circ R(t, \theta)$$

.

■

Remarque 6.1.9. En utilisant la relation $R(t, \theta) = (R(\theta, t))^{-1}$, on voit que, si $A(\theta)$ est inversible, alors $R'_t(t, \theta)$ l'est aussi et $R(t, \theta)$ est dérivable par rapport à θ , de dérivée :

$$R'_\theta(t, \theta) = \frac{\partial}{\partial \theta} (R(\theta, t))^{-1} = \left(\frac{\partial}{\partial \theta} R(\theta, t) \right)^{-1} = (R(\theta, t))^{-1} \circ A(\theta)^{-1}$$

Remarque 6.1.10. L'équation différentielle vérifiée par la résolvante est d'une part plus compliquée que l'équation initiale vérifiée par la fonction car elle est à valeurs dans $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ et non pas dans \mathbb{R}^m mais d'autre part, cela peut aussi se révéler être un avantage car on peut exploiter la structure d'algèbre de Banach unitaire de $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$.

En général, on ne sait pas expliciter la résolvante d'une équation différentielle linéaire sauf dans certains cas particuliers où la matrice $A(t)$ a des propriétés particulières. Avant d'étudier ces cas simples, rappelons les propriétés des exponentielles de matrices que l'on a déjà rencontrées au chapitre précédent, cf. 5.2.11 :

Définition 6.1.11. Soit $A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$. L'exponentielle de A est définie par

$$e^A = \sum_{n=0}^{+\infty} \frac{A^n}{n!}$$

où A^n est la n -ième puissance de composition de A par elle-même.

Remarque 6.1.12. La série ci-dessus est normalement convergente donc convergente. En effet, on peut écrire

$$\left\| \frac{A^n}{n!} \right\| \leq \frac{1}{n!} \|A\|^n$$

et la série numérique $\sum \frac{1}{n!} \|A\|^n$ est convergente.

Théorème 6.1.13. (Propriétés de l'exponentielle)

Si A et $B \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ commutent, on a

$$e^{A+B} = e^A e^B$$

DÉMONSTRATION : C'est la même que dans le cas scalaire : comme les séries définissant e^A et e^B sont normalement convergentes, on a le droit de sommer comme on veut, d'où :

$$\begin{aligned} e^A \circ e^B &= \sum_{m,n=0}^{+\infty} \frac{A^m B^n}{n! m!} = \sum_{p=0}^{+\infty} \sum_{k=0}^p \frac{A^k B^{p-k}}{k! (p-k)!} = \\ &= \sum_{p=0}^{+\infty} \frac{1}{p!} \sum_{k=0}^p C_p^k A^k B^{p-k} = \sum_{p=0}^{+\infty} \frac{1}{p!} (A+B)^p = e^{A+B} \end{aligned}$$

■

Calcul de e^A . Rappelons que l'on peut exprimer ces résultats à l'aide des valeurs propres de A :

Soient $\{\lambda_i \mid i = 1, 2, \dots, k\}$ les valeurs propres de A , $\{r_i \mid i = 1, 2, \dots, k\}$ leur ordre de multiplicité dans le polynôme caractéristique de A et $\{\pi_i \mid i = 1, 2, \dots, k\}$ les projecteurs spectraux sur les sous espaces spectraux $E_i = \ker(A - \lambda_i I)^{r_i}$, de dimension r_i .

Alors, il existe des endomorphismes nilpotents $\{N_i \mid i = 1, 2, \dots, k\}$ dans $\mathcal{L}(E_i; E_i)$, d'indice k_i tels que :

$$A = \sum_{i=1}^k (\lambda_i I_{E_i} + N_i) \pi_i$$

$$e^{tA} = \sum_{i=1}^k e^{\lambda_i t} e^{tN_i} \pi_i = \sum_{i=1}^k e^{\lambda_i t} (I_{E_i} + tN_i + \dots + \frac{t^{k_i-1}}{(k_i-1)!} N_i^{k_i-1}) \pi_i$$

■

Cas Particulier. 6.1.14. Soit EDO $X'(t) = A(t)X(t)$. Si pour tous $t, u \in I$,

$$A(t)A(u) = A(u)A(t)$$

alors

$$R(t, t_0) = \exp \left(\int_{t_0}^t A(s) ds \right)$$

DÉMONSTRATION : Posons $M(t) = \exp \left(\int_{t_0}^t A(s) ds \right) \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ et remarquons que $M(t_0) = I_{\mathbb{R}^m}$ et que $M(t)$ commute avec $A(t)$. L'hypothèse de commutation entraîne de plus que $\int_{t_1}^{t_2} A(s) ds$ et $\int_{t_3}^{t_4} A(s) ds$ commutent pour tous $t_1, t_2, t_3, t_4 \in I$.

Donc :

$$M(t+h) = \exp \left(\int_{t_0}^t A(s) ds + \int_t^{t+h} A(s) ds \right) = M(t) \exp \left(\int_t^{t+h} A(s) ds \right)$$

Or, on a : $\int_t^{t+h} A(s) ds = hA(t) + h\varepsilon_t(h)$ et en utilisant le développement en série de l'exponentielle :

$$M(t+h) = M(t)(Id + hA(t) + h\varepsilon_t(h)) = M(t) + hM(t)A(t) + h\varepsilon_t^1(h)$$

ce qui montre que $M'(t) = M(t)A(t) = A(t)M(t)$. M vérifie l'équation caractéristique de la résolvante dans $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$ et donc est la résolvante de EDO. Comme $M(t_0) = I_{\mathbb{R}^m}$, on a donc bien $M(t) = R(t, t_0)$. ■

Exemples 6.1.15.

i) Si A et B sont des matrices qui commutent et $A(t) = f(t)A + g(t)B$ où f et g sont des fonctions scalaires, alors

$$R(t, t_0) = \exp \left(\int_{t_0}^t f(s) ds A + \int_{t_0}^t g(s) ds B \right) = \exp \left(\int_{t_0}^t f(s) ds A \right) \exp \left(\int_{t_0}^t g(s) ds B \right)$$

ii) Le système

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 1/t & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

défini sur \mathbb{R}_*^+ ne vérifie pas l'hypothèse de commutation du Cas Particulier 6.1.14 et on a :

$$R(t, t_0) \neq \exp \left(\int_{t_0}^t A(s) ds \right)$$

DÉMONSTRATION : On calcule la résolvante en résolvant le système différentiel triangulaire :

$$\begin{cases} x'(t) = \frac{1}{t}x(t) + ty(t) \\ y'(t) = y(t) \end{cases}$$

et on trouve

$$R(t, t_0) = \begin{pmatrix} \frac{t}{t_0} & t(e^{t-t_0} - 1) \\ 0 & e^{t-t_0} \end{pmatrix}$$

Pour calculer l'exponentielle de

$$B(t) = \int_{t_0}^t A(s) ds = \begin{pmatrix} \ln \frac{t}{t_0} & \frac{t^2 - t_0^2}{2} \\ 0 & t - t_0 \end{pmatrix}$$

on se place dans une base de vecteurs propres de $B(t)$, la matrice de passage est alors

$$P(t) = \begin{pmatrix} 1 & -\frac{t^2 - t_0^2}{2} \\ 0 & \ln \frac{t}{t_0} - (t - t_0) \end{pmatrix}$$

et on trouve

$$e^{B(t)} = P(t) \begin{pmatrix} \frac{t}{t_0} & 0 \\ 0 & e^{t-t_0} \end{pmatrix} P^{-1}(t) = \begin{pmatrix} \frac{t}{t_0} & \varphi(t) \\ 0 & e^{t-t_0} \end{pmatrix}$$

avec

$$\varphi(t) = \frac{t^2 - t_0^2}{2(\ln \frac{t}{t_0} - (t - t_0))} \left(\frac{t}{t_0} - e^{t-t_0} \right)$$

■

6.2 Equations différentielles linéaires avec deuxième membre

Définition 6.2.1. On appelle *équation différentielle linéaire*, définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R}^m , avec *deuxième membre* une équation différentielle de la forme

$$X'(t) = A(t)X(t) + B(t) \text{ avec } \forall t \in I, A(t) \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m) \text{ et } B \in \mathcal{C}(I; \mathbb{R}^m)$$

Théorème 6.2.2. (Solution d'une équation différentielle linéaire avec deuxième membre)

Si l'application $t \rightarrow A(t)$ est continue de I dans $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$, l'équation différentielle linéaire avec deuxième membre :

$$EDO \quad X'(t) = A(t)X(t) + B(t)$$

admet une et une seule solution, définie sur I et vérifiant la condition initiale : $X(t_0) = x_0$. Cette solution est donnée par la formule

$$X(t) = R(t, t_0)x_0 + \int_{t_0}^t R(t, \theta)B(\theta) d\theta$$

DÉMONSTRATION : La démonstration de l'existence et l'unicité de la solution est la même que celle des équations différentielles linéaires sans deuxième membre (6.1.2)

Pour obtenir la forme de la solution, on utilise la **méthode de variation de la constante** : On pose

$$X(t) = R(t, t_0)Y(t)$$

D'où

$$X'(t) = R'_t(t, t_0)Y(t) + R(t, t_0)Y'(t)$$

C'est à dire

$$X'(t) = A(t)R(t, t_0)Y(t) + R(t, t_0)Y'(t)$$

En reportant dans l'équation différentielle, il vient :

$$R(t, t_0)Y'(t) = B(t)$$

C'est à dire

$$Y'(t) = R(t_0, t)B(t)$$

La condition initiale est $Y(t_0) = x_0$, d'où

$$Y(t) = x_0 + \int_{t_0}^t R(t_0, \theta)B(\theta) d\theta$$

et en revenant à X :

$$X(t) = R(t, t_0)x_0 + R(t, t_0) \int_{t_0}^t R(t_0, \theta)B(\theta) d\theta =$$

$$R(t, t_0)x_0 + \int_{t_0}^t R(t, t_0)R(t_0, \theta)B(\theta) d\theta =$$

$$R(t, t_0)x_0 + \int_{t_0}^t R(t, \theta)B(\theta) d\theta$$

■

Corollaire 6.2.3. L'ensemble des solutions d'une équation différentielle linéaire avec deuxième membre sur \mathbb{R}^m est un espace affine dont l'espace vectoriel associé est isomorphe à \mathbb{R}^m .

DÉMONSTRATION : La partie linéaire correspond à l'application $x_0 \rightarrow R(t, t_0)x_0$ et la partie affine à la fonction particulière $\int_{t_0}^t R(t, \theta)B(\theta) d\theta$. ■

6.3 Equations différentielles linéaires à coefficients constants

Théorème 6.3.1. (Résolvante d'une équation différentielle linéaire à coefficients constants)

La résolvante de l'équation différentielle linéaire à coefficients constants

$$EDO X'(t) = AX(t) , \text{ avec } A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^m)$$

est

$$R(t_1, t_2) = e^{(t_1 - t_2)A}$$

DÉMONSTRATION : C'est une conséquence du Cas Particulier 6.1.14, mais on peut en donner une démonstration directe : comme la série définissant e^{tA} est normalement convergente, on a le droit de dériver terme à terme. D'où

$$\frac{d}{dt}(e^{tA}) = \sum_{n=1}^{+\infty} \frac{t^{n-1} A^n}{(n-1)!}$$

Donc aussi

$$\frac{d}{dt}(e^{tA}) = Ae^{tA}$$

Si l'on pose $R(t, \theta) = e^{(t-\theta)A}$, on aura :

$$R'_t(t, \theta) = Ae^{(t-\theta)A} = AR(t, \theta) \text{ et } R(\theta, \theta) = I_{\mathbb{R}^m}$$

ce qui montre que R ainsi défini est bien la résolvante de l'équation différentielle linéaire à coefficients constants considérée. ■

Corollaire 6.3.2. La solution de l'équation différentielle linéaire à coefficients constants avec deuxième membre

$$EDO X'(t) = AX(t) + B(t)$$

telle que $X(t_0) = x_0$ est :

$$X(t) = e^{(t-t_0)A} x_0 + \int_{t_0}^t e^{(t-\theta)A} B(\theta) d\theta$$

6.4 Equations différentielles linéaires d'ordre m à coefficients constants

On considère l'équation différentielle d'ordre m :

$$EDO_m \ a_m y^{(m)}(t) + \dots + a_1 y'(t) + a_0 y(t) = b(t)$$

où y est la fonction inconnue définie sur $I \in \mathbb{R}$, à valeurs dans \mathbb{R} , b est une fonction définie sur I à valeurs dans \mathbb{R} et les coefficients a_i , $i = 1, \dots, m$ sont des constantes appartenant à \mathbb{R} , avec $a_m \neq 0$.

Conformément à la théorie générale, cf. 5.1.13, on va remplacer cette équation d'ordre m dans \mathbb{R} par une équation d'ordre 1 dans \mathbb{R}^m . Pour cela, posons

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & & & & & \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & 0 & 1 \\ c_0 & c_1 & c_2 & \dots & c_{m-2} & c_{m-1} \end{pmatrix}, \quad c_j = -\frac{a_j}{a_m}, \quad B(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{1}{a_m} b(t) \end{pmatrix}$$

alors EDO_m est équivalente au système :

$$S_m \ Y'(t) = AY(t) + B(t)$$

où l'on a posé

$$Y(t) = \begin{pmatrix} y(t) \\ y'(t) \\ y''(t) \\ \vdots \\ y^{(m-1)}(t) \end{pmatrix}$$

Pour appliquer les résultats généraux d'existence et unicité des solutions d'équations différentielles linéaires dans \mathbb{R}^m , on doit donner comme condition initiale de EDO_m pour $t_0 \in I$ et $X_0 \in \mathbb{R}^m$, conforme à la théorie générale, cf. 5.1.14, c'est à dire :

$$Y(t_0) = \begin{pmatrix} y(t_0) \\ y'(t_0) \\ y''(t_0) \\ \vdots \\ y^{(m-1)}(t_0) \end{pmatrix} = X_0$$

Dans ces conditions, on obtient alors comme conséquence du Corollaire 6.2.3 :

Théorème 6.4.1. *L'ensemble des solutions de l'équation différentielle EDO_m est un espace affine dont l'espace vectoriel associé est de dimension m .*

DÉMONSTRATION : Il est clair qu'à chaque solution y de EDO_m correspond une et une seule solution Y de S_m . La relation entre les solutions de EDO_m et S_m est donc bijective. Comme cette correspondance est évidemment linéaire, on a donc une correspondance linéaire bijective entre les solutions de EDO_m et de S_m , ce qui prouve le théorème, en appliquant le Corollaire 6.2.3 à S_m . ■

Maintenant que l'on sait quelle est la structure des solutions de EDO_m grâce à l'étude du système S_m et à l'existence de la résolvante, on peut employer des méthodes plus artisanales pour trouver ces solutions : il s'agit d'exhiber m solutions indépendantes de l'équation sans second membre $EDO_m|_0$ et une solution particulière de l'équation EDO_m .

Résolution de l'équation sans second membre $EDO_m|_0$.

On se place sur le corps \mathbb{C} , les solutions réelles s'obtiennent simplement en prenant la partie réelle et la partie imaginaire des solutions complexes.

Remarque 6.4.2. *Le polynôme caractéristique de la matrice A est*

$$P(\lambda) = (-1)^{m+1}[a_m\lambda^m + \dots a_1\lambda + a_0]$$

On cherche des solutions de la forme :

$$y(t) = e^{\lambda t}, \lambda \in \mathbb{C}$$

Comme $y^{(j)}(t) = \lambda^j e^{\lambda t}$, on voit que y est solution de $EDO_m|_0$ si et seulement si λ est racine du polynôme caractéristique

$$P(\lambda) = (-1)^{m+1}[a_m\lambda^m + \dots a_1\lambda + a_0]$$

• Si P n'a que des racines simples, ce polynôme possède m racines distinctes $\lambda_1, \dots, \lambda_m$ et on obtient m solutions distinctes

$$t \rightarrow e^{\lambda_j t}, 1 \leq j \leq m.$$

On verra plus loin que ces solutions sont linéairement indépendantes sur \mathbb{C} . L'ensemble des solutions est donc l'espace vectoriel de dimension m des fonctions

$$y(t) = \alpha_1 e^{\lambda_1 t} + \dots + \alpha_m e^{\lambda_m t}, \alpha_j \in \mathbb{C}.$$

• Si P a des racines multiples, on peut écrire :

$$P(\lambda) = a_m \prod_{j=1}^s (\lambda - \lambda_j)^{m_j}$$

où m_j est l'ordre de multiplicité de la racine λ_j . On a donc :

$$m_1 + \dots + m_s = m.$$

Considérons l'opérateur différentiel

$$P\left(\frac{d}{dt}\right) = \sum_{i=0}^m a_i \frac{d^i}{dt^i}$$

L'équation différentielle $EDO_m|_0$ peut se ré-écrire

$$P\left(\frac{d}{dt}\right)y(t) = 0$$

et d'autre part, on a les formules :

$$\forall \lambda \in \mathbb{C}, \forall \lambda \in \mathbb{C}, P\left(\frac{d}{dt}\right)e^{\lambda t} = P(\lambda)e^{\lambda t}, \frac{d^q}{d\lambda^q}e^{\lambda t} = t^q e^{\lambda t}$$

Comme les dérivées partielles $\frac{d}{dt}$ et $\frac{d}{d\lambda}$ commutent d'après le théorème de Schwarz, on obtient alors :

$$P\left(\frac{d}{dt}\right)(t^q e^{\lambda t}) = P\left(\frac{d}{dt}\right)\left(\frac{d^q}{d\lambda^q}e^{\lambda t}\right) = \frac{d^q}{d\lambda^q}\left(P\left(\frac{d}{dt}\right)e^{\lambda t}\right) = \frac{d^q}{d\lambda^q}\left(P(\lambda)e^{\lambda t}\right)$$

d'où grâce à la formule de Leibnitz

$$P\left(\frac{d}{dt}\right)(t^q e^{\lambda t}) = \sum_{i=0}^q C_q^i P^{(i)}(\lambda) e^{\lambda t} t^{q-i}.$$

Comme λ_j est racine de multiplicité d'ordre m_j de P , on a $P^{(i)}(\lambda_j) = 0$ pour $0 \leq i \leq m_j - 1$, et $P^{(m_j)}(\lambda_j) \neq 0$. On en déduit

$$P\left(\frac{d}{dt}\right)(t^q e^{\lambda_j t}) = 0, \quad 0 \leq q \leq m_j - 1.$$

L'équation différentielle $EDO_p|_0$ admet donc les solutions

$$y(t) = t^q e^{\lambda_j t}, \quad 0 \leq q \leq m_j - 1, \quad 1 \leq j \leq s$$

soit au total $m_1 + \dots + m_s = m$ solutions.

Il reste à voir que dans les deux cas, P ayant des racines simples ou des racines multiples, les solutions que nous venons de trouver sont indépendantes et que par conséquent nous les avons toutes obtenues :

Lemme 6.4.3. *Si $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ sont des nombres complexes deux à deux distincts, alors les fonctions :*

$$y_{j,q}(t) = t^q e^{\lambda_j t}, \quad 1 \leq j \leq s, \quad q \in \mathbb{N}$$

sont linéairement indépendantes.

DÉMONSTRATION : On va procéder par l'absurde : considérons une combinaison linéaire finie

$$\sum \alpha_{j,q} y_{j,q} = 0, \quad \alpha_{j,q} \in \mathbb{C}$$

Si les coefficients sont non tous nuls, soit N le maximum des entiers q tels qu'il existe j avec $\alpha_{j,q} \neq 0$. Supposons par exemple $\alpha_{1,N} \neq 0$. On pose alors

$$Q(\lambda) = (\lambda - \lambda_1)^N (\lambda - \lambda_2)^{N+1} \dots (\lambda - \lambda_s)^{N+1}$$

Il vient $Q^{(i)}(\lambda_j) = 0$ pour $j \geq 2$ et $0 \leq i \leq N$, tandis que $Q^{(i)}(\lambda_1) = 0$ pour $0 \leq i < N$ et $Q^{(N)}(\lambda_1) \neq 0$. On en déduit, en utilisant les formules précédentes :

$$Q\left(\frac{d}{dt}\right)(t^q e^{\lambda_j t}) = \sum_{i=0}^q C_q^i Q^{(i)}(\lambda_j) t^{q-i} e^{\lambda_j t} = 0$$

pour $0 \leq q \leq N$ et $1 \leq j \leq s$, sauf si $q = N$ et $j = 1$ auquel cas

$$Q\left(\frac{d}{dt}\right)(t^N e^{\lambda_1 t}) = Q^{(N)}(\lambda_1) e^{\lambda_1 t}.$$

En appliquant l'opérateur $Q\left(\frac{d}{dt}\right)$ à la relation $\sum \alpha_{j,q} t^q e^{\lambda_j t} = 0$ on obtient alors : $\alpha_{1,N} Q^{(N)}(\lambda_1) e^{\lambda_1 t} = 0$, ce qui est absurde puisque $\alpha_{1,N} \neq 0$ et $Q^{(N)}(\lambda_1) \neq 0$.

On obtient donc une contradiction et ceci prouve que les coefficients $\alpha_{j,q}$ sont tous nuls et par suite que les fonctions $y_{j,q}$, $1 \leq j \leq s$, $q \in \mathbb{N}$ sont linéairement indépendantes. ■

On peut donc énoncer :

Théorème 6.4.4. *Lorsque le polynôme caractéristique $P(\lambda)$ a des racines complexes $\lambda_1, \dots, \lambda_s$ d'ordre de multiplicité respectif m_1, \dots, m_s , l'ensemble des solutions de l'équation différentielle $EDO_m|_0$ est le \mathbb{C} -espace vectoriel de dimension m ayant pour base les fonctions*

$$t \rightarrow t^q e^{\lambda_j t}, \quad 1 \leq j \leq s, \quad 0 \leq q \leq m_j - 1.$$

Résolution de l'équation avec second membre EDO_m .

On utilise une méthode analogue à la **méthode de variation de la constante**, définie au paragraphe 2 :

Soit v_1, v_2, \dots, v_m une base des solutions de l'équation sans second membre $EDO_m|_0$. On cherche une solution particulière par la méthode de variation de la constante, appliquée au système différentiel d'ordre 1 dans \mathbb{R}^m ,

$$S_m Y'(t) = AY(t) + B(t).$$

Le système sans second membre $S_m|_0 Y'(t) = AY(t)$ admet pour base de solutions les fonctions :

$$V_1 = \begin{pmatrix} v_1 \\ v_1' \\ \vdots \\ v_1^{(m-1)} \end{pmatrix}, \quad V_2 = \begin{pmatrix} v_2 \\ v_2' \\ \vdots \\ v_2^{(m-1)} \end{pmatrix}, \quad \dots, \quad V_m = \begin{pmatrix} v_m \\ v_m' \\ \vdots \\ v_m^{(m-1)} \end{pmatrix}.$$

On cherche alors une solution particulière de S_m sous la forme

$$Y(t) = \alpha_1(t)V_1(t) + \dots + \alpha_m(t)V_m(t)$$

Comme $V_j' = AV_j$, il vient

$$Y'(t) = \sum_{j=1}^m \alpha_j(t)V_j'(t) + \sum_{j=1}^m \alpha_j'(t)V_j(t) = AY(t) + \sum_{j=1}^m \alpha_j'(t)V_j(t)$$

Il suffit donc de choisir les α_j tels que $\sum \alpha_j'(t)V_j(t) = B(t)$, c'est à dire :

$$\begin{cases} \alpha_1'(t)v_1(t) + \dots + \alpha_m'(t)v_m(t) = 0 \\ \dots \\ \alpha_1'(t)v_1^{(m-2)}(t) + \dots + \alpha_m'(t)v_m^{(m-2)}(t) = 0 \\ \alpha_1'(t)v_1^{(m-1)}(t) + \dots + \alpha_m'(t)v_m^{(m-1)}(t) = \frac{1}{a_m}b(t) \end{cases}$$

On obtient ainsi un système linéaire de m équations par rapport aux m inconnues $\alpha_1'(t), \dots, \alpha_m'(t)$. Le déterminant de ce système est non nul pour tout $t \in \mathbb{R}$ car les vecteurs $V_1(t), \dots, V_m(t)$ sont linéairement indépendants : en effet, si une combinaison linéaire $Y = \beta_1 V_1 + \dots + \beta_m V_m$ est nulle pour un $t \in \mathbb{R}$, par le théorème d'unicité, on a $Y \equiv 0$, donc $\beta_1 = \dots = \beta_m = 0$.

La résolution de ce système permet de calculer $\alpha_1', \dots, \alpha_m'$ puis $\alpha_1, \dots, \alpha_m$ par intégration, d'où la solution particulière cherchée :

$$y(t) = \alpha_1(t)v_1(t) + \dots + \alpha_m(t)v_m(t).$$

Exemple 6.4.5. Résoudre $E \ y''(t) + 4y(t) = \operatorname{tg} t$ avec $t \in]-\pi/2, \pi/2[$ et déterminer la résolvante $R(t, 0)$ du système d'ordre 2 associé.

La solution générale de l'équation sans second membre est $y(t) = a \cos 2t + b \sin 2t$. En posant $z = y'$, le système d'ordre 2 associé à E est

$$\begin{pmatrix} y'(t) \\ z'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$$

Pour trouver la résolvante, il faut introduire la condition initiale $y(0) = x_0, y'(0) = x_1$, c'est à dire : $a = x_0, 2b = x_1$. D'où

$$R(t, 0) = \begin{pmatrix} \cos 2t & \frac{\sin 2t}{2} \\ -2 \sin 2t & \cos 2t \end{pmatrix}$$

On cherche une solution particulière de l'équation avec second membre par la méthode de variation des constantes, c'est à dire qu'on résoud le système :

$$\begin{cases} a'(t) \cos 2t + b'(t) \sin 2t = 0 \\ -2a'(t) \sin 2t + 2b'(t) \cos 2t = \operatorname{tg} t \end{cases}$$

soit

$$\begin{cases} a'(t) = -\frac{1}{2} \operatorname{tg} t \sin 2t = -\sin^2 t = \frac{1 - \cos 2t}{2} \\ b'(t) = \frac{1}{2} \operatorname{tg} t \cos 2t = \sin t \cos t - \frac{\operatorname{tg} t}{2} \end{cases}$$

D'où

$$\begin{cases} a(t) = a + \frac{t}{2} - \frac{\sin 2t}{4} \\ b(t) = b + \frac{\sin^2 t}{2} + \frac{1}{2} \ln \cos t \end{cases}$$

La solution générale de E est donc $y(t) = (a + \frac{t}{2} - \frac{\sin 2t}{4}) \cos 2t + (b + \frac{\sin^2 t}{2} + \frac{1}{2} \ln \cos t) \sin 2t$.

6.5 Wronskien d'un système de solutions

On considère l'équation différentielle

$$EDO \quad Y'(t) = A(t)Y(t)$$

où les fonctions sont à valeurs dans l'espace \mathbb{R}^m .

Définition 6.5.1. Le Wronskien d'un système de m solutions (Y_1, \dots, Y_m) de EDO est

$$W(t) = \det(Y_1(t), \dots, Y_m(t))$$

Théorème 6.5.2. Si $V_j = Y_j(t_0)$ pour $j = 1, 2, \dots, m$, alors $Y_j(t) = R(t, t_0)V_j$ et on a

$$W(t) = \exp\left(\int_{t_0}^t \operatorname{tr}(A(s)) ds\right) \det(V_1, \dots, V_m).$$

DÉMONSTRATION : On a besoin d'un lemme :

Lemme 6.5.3. Si $A = (a_{ij})$ est une matrice $m \times m$ et $h \in \mathbb{R}$, alors

$$\det(I_m + hA) = 1 + \alpha_1 h + \dots + \alpha_m h^m$$

avec $\alpha_1 = \operatorname{tr} A = \sum_{i=1}^m a_{ii}$.

En effet, par définition, $\det(I_m + hA)$ est un polynôme de degré m en h . Le terme diagonal est

$$(1 + ha_{11}) \dots (1 + ha_{mm}) = 1 + h \sum_{i=1}^m a_{ii} + h^2 \dots$$

et les termes non diagonaux sont multiples de h^2 . ■

Revenons à la démonstration du théorème : on peut écrire

$$W(t) = \det(R(t, t_0)) \det(V_1, \dots, V_m).$$

On est donc amené à calculer la quantité $\Delta(t) = \det(R(t, t_0))$. Pour cela, on va montrer que $\Delta(t)$ vérifie une équation différentielle simple. On a :

$$\Delta(t+h) = \det(R(t+h, t_0)) = \det(R(t+h, t)R(t, t_0)) = \det(R(t+h, t))\Delta(t).$$

Comme $R(t, t) = I_m$ et $\frac{d}{du}R(u, t)|_{u=t} = A(t)R(t, t) = A(t)$, la formule de Taylor donne

$$R(t+h, t) = I_m + hA(t) + o(h)$$

$$\det(R(t+h, t)) = \det(I_m + hA(t)) + o(h).$$

Le lemme entraîne alors

$$\det(R(t+h, t)) = 1 + h\operatorname{tr}(A(t)) + o(h)$$

$$\Delta(t+h) = \Delta(t) + h\operatorname{tr}(A(t))\Delta(t) + o(h)$$

On en déduit

$$\Delta'(t) = \operatorname{tr}(A(t))\Delta(t)$$

et comme $\Delta(t_0) = \det(R(t_0, t_0)) = \det I_m = 1$, il vient :

$$\det(R(t, t_0)) = \Delta(t) = \exp\left(\int_{t_0}^t \operatorname{tr}A(u) du\right)$$

$$W(t) = \exp\left(\int_{t_0}^t \operatorname{tr}A(u) du\right) \det(V_1, \dots, V_m).$$

■

7

MÉTHODES D'APPROXIMATION NUMÉRIQUE À UN PAS DES ÉQUATIONS DIFFÉRENTIELLES.

Dans ce chapitre, on cherche à trouver des méthodes permettant de résoudre numériquement une *EDO* avec condition initiale :

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

où f est une fonction continue sur $[t_0, t_0 + T] \times \mathbb{R}^m$, à valeurs dans \mathbb{R}^m . Dans tout ce chapitre, on supposera que f est globalement Lipschitzienne par rapport à x , uniformément en $t \in [0, T]$, avec constante L . Ceci assure que l'*EDO* $y'(t) = f(t, y(t))$ avec condition initiale $y(t_0) = y_0$ admet une unique solution sur $[0, T]$.

Etant donnée une subdivision $t_0 < t_1 < \dots < t_N = t_0 + T$ de l'intervalle $[t_0, t_0 + T]$, on cherche à déterminer des valeurs approchées y_0, y_1, \dots, y_N des valeurs $y(t_0), y(t_1), \dots, y(t_N)$ prises par la solution exacte y aux temps successifs t_0, t_1, \dots, t_N . On notera les pas successifs

$$h_n = t_{n+1} - t_n, \quad 0 \leq n \leq N - 1$$

et

$$h_{max} = \max_n h_n.$$

Définition 7.0.1. On appelle méthode à un pas, une méthode permettant de calculer y_{n+1} à partir de y_n .

7.1 Convergence des schémas numériques explicites

Dans cette section, on va étudier la famille dite des schémas explicites. On se donne une fonction F de trois variables $(t, x, h) \in [0, T] \times \mathbb{R}^m \times [0, T]$, à valeurs dans \mathbb{R}^m que l'on suppose continue. On utilisera les deux propriétés suivantes, dites de consistance et de stabilité :

Hypothèses. 7.1.1. 1) On dit que F vérifie l'hypothèse de consistance si :

$$F(t, x, 0) = f(t, x), \quad \forall t \in [0, T], \quad \forall x \in \mathbb{R}^m$$

2) On dit que F vérifie l'hypothèse de stabilité s'il existe une constante $L_1 > 0$ telle que

$$|F(t, b, h) - F(t, c, h)| \leq L_1 |b - c| \quad \forall t \in [0, T], \quad \forall b, c \in \mathbb{R}^m, \quad \forall h \in [0, T].$$

On définit alors le schéma explicite par :

Définition 7.1.2. Soit $0 < h \leq T$ un pas de temps. Le schéma explicite associé à F est donné par la récurrence suivante :

$$y_{n+1} = y_n + hF(nh, y_n, h), \quad n = 0, 1, 2, \dots, \quad y_0 = a,$$

où a est la donnée initiale et tant que $nh \leq T$.

Un tel schéma est appelé explicite parce que la valeur de y_{n+1} ne nécessite pas la résolution d'une équation : elle est donnée explicitement en fonction de nh , y_n et f .

Exemple 7.1.3. Le schéma d'Euler, défini au chapitre 5, 5.1.15 correspond au cas très particulier où F ne dépend pas de h et est donné par :

$$F(t, x, h) = f(t, x).$$

Le schéma numérique d'Euler est donc donné par la formule :

$$y_{n+1} = y_n + hf(nh, y_n)$$

Mais, il est judicieux, comme on le verra plus loin, de considérer des fonctions plus compliquées, par exemple :

Exemple 7.1.4. Le schéma explicite dit du point milieu (ou encore d'Euler modifié) est défini par la fonction F :

$$F(t, x, h) = f(t + h/2, x + hf(t, x)/2).$$

Le schéma numérique du point milieu est donc donné par la formule :

$$y_{n+1} = y_n + hf\left(nh + \frac{h}{2}, y_n + \frac{h}{2}f(nh, y_n)\right)$$

Dans les deux cas, les hypothèses de stabilité et consistance sont bien satisfaites. Pour la méthode d'Euler, c'est évident. Montrons la stabilité du schéma du point milieu :

Lemme 7.1.5. *Le schéma du point milieu est stable, au sens de la stabilité définie en 7.1.1 avec la constante :*

$$L_1 = L(1 + LT/2).$$

DÉMONSTRATION : Soit b, c fixés dans \mathbb{R}^m et t dans $[0, T]$. On a

$$\begin{aligned} |F(t, b, h) - F(t, c, h)| &= |f(t + h/2, b + h/2f(t, b)) - f(t + h/2, c + h/2f(t, c))| \\ &\leq L|b + h/2f(t, b) - c - h/2f(t, c)| \leq L(|b - c| + h/2|f(t, b) - f(t, c)|) \\ &\leq L(1 + Lh/2)|b - c| \leq L(1 + LT/2)|b - c|. \end{aligned}$$

■

La suite de valeurs $(y_n), n = 0, 1, 2, \dots$ obtenue par un schéma numérique avec $y_0 = y(0)$ dépend de h , on peut la noter plus lourdement $y_{n,h}$. Elle permet pour chaque $h > 0$ de construire une fonction $y_h(t)$ affine par morceaux définie pour $t \in I = [0, T[$ et obtenue en interpolant les y_n aux instants $t = nh$, c'est à dire :

$$y_h(t) = y_n + (t - nh)F(nh, y_n, h) \text{ pour } nh \leq t \leq (n + 1)h$$

Une façon équivalente de définir cette fonction est de noter pour $t \geq 0$, h fois la partie entière de t/h par t_h , et de poser,

$$y_h(t) = y_h(t_h) + (t - t_h)F(t_h, y_h(t_h), h).$$

On s'attend à ce que y_h soit une approximation de la solution y , convergeant vers y lorsque le pas de temps h_{max} tend vers 0. Pour mesurer cette convergence, on évaluera l'erreur au n -ième pas de temps :

Définition 7.1.6. *L'erreur au n -ième pas de temps est égale à la différence entre la valeur prise par la solution y au point nh et celle prise par l'approximation y_h , c'est à dire*

$$e_n = y_h(nh) - y(nh) = y_n - y(nh)$$

Définition 7.1.7. *On dira que le schéma est convergent si*

$$\sup_{nh \leq T} |e_n| \rightarrow 0, \text{ quand } h \rightarrow 0$$

Pour évaluer l'erreur, on montre d'abord :

Proposition 7.1.8. *Soit $F : (t, x, h) \in [0, T] \times \mathbb{R}^m \times [0, T] \rightarrow F(t, x, h) \in \mathbb{R}^m$ continue. Alors la propriété de consistance, 7.1.1, entraîne que pour toute fonction $t \in [0, T] \rightarrow z(t) \in \mathbb{R}^m$ continue, et tout $h \in]0, T]$, le nombre*

$$\eta(z, h) = \sup_{0 \leq t \leq T-h} |F(t, z(t), h) - \frac{1}{h} \int_t^{t+h} f(s, z(s)) ds|$$

tend vers 0 lorsque $h \rightarrow 0$.

DÉMONSTRATION : Fixons z et posons

$$G(t, h) = F(t, z(t), h), \quad \forall t \in [0, T], \quad \forall h \in [0, T],$$

ce qui définit une fonction continue sur le compact $[0, T] \times [0, T]$. Elle est donc uniformément continue. Grace à l'hypothèse de consistance, on peut réécrire

$$\begin{aligned} F(t, z(t), h) - \frac{1}{h} \int_t^{t+h} f(s, z(s)) ds &= \frac{1}{h} \int_t^{t+h} (F(t, z(t), h) - F(s, z(s), 0)) ds \\ &= \frac{1}{h} \int_t^{t+h} (G(t, h) - G(s, 0)) ds. \end{aligned}$$

On a donc, en utilisant que la norme de l'intégrale est plus petite que l'intégrale de la norme :

$$\begin{aligned} |F(t, z(t), h) - \frac{1}{h} \int_t^{t+h} f(s, z(s)) ds| &\leq \frac{1}{h} \int_t^{t+h} |G(t, h) - G(s, 0)| ds \\ &\leq \sup_{t \leq s \leq t+h} |G(t, h) - G(s, 0)|. \end{aligned}$$

Ainsi, on a obtenu

$$\eta(z, h) \leq \sup_{0 \leq t \leq s \leq t+h \leq T} |G(t, h) - G(s, 0)|.$$

Comme G est uniformément continue, le terme de droite tend vers 0 avec h ce qui termine la démonstration de la Proposition 7.1.8. ■

On peut montrer qu'en fait, du moment que F est continue, la propriété de consistance est non seulement suffisante mais aussi nécessaire pour obtenir le résultat de la Proposition 7.1.8 : il suffit considérer le cas où z est une fonction constante.

Remarque importante. Dans le cas particulier où pour z , on prend une solution y de l'EDO, on a

$$\frac{1}{h} \int_t^{t+h} f(s, y(s)) ds = \frac{1}{h} \int_t^{t+h} y'(s) ds = \frac{y(t+h) - y(t)}{h}$$

et on a donc

$$\eta(y, h) = \sup_{0 \leq t \leq T-h} |F(t, y(t), h) - \frac{y(t+h) - y(t)}{h}|.$$

Définition 7.1.9. Le nombre $\eta(y, h)$ défini ci-dessus est appelé erreur de consistance du schéma.

A l'aide de la proposition 7.1.8, on prouve ensuite le Théorème de convergence des schémas explicites :

Théorème 7.1.10. Soit y l'unique solution de l'EDO $y'(t) = f(t, y(t))$ telle que $y(0) = a$ et soit la solution approchée y_n définie par le schéma explicite : $y_{n+1} = y_n + hF(nh, y_n, h)$, $n = 0, 1, 2, \dots$, $y_0 = a$. On suppose que F est continue et vérifie les hypothèses de consistance et de stabilité. Alors on a convergence du schéma, avec l'estimation d'erreur

$$\sup_{nh \leq T} |e_n| \leq \sup_{nh \leq T} |y(nh) - y_n| \leq M\eta(y, h)$$

où

$$M = \frac{\exp(L_1 T) - 1}{L_1}.$$

et L_1 est la constante de stabilité de F .

Remarques. On a donc en particulier la convergence du schéma d'Euler et du schéma du point milieu.

Définition 7.1.11. Le nombre M s'appelle constante de stabilité du schéma.

Dans la constante de stabilité du schéma M , l'exponentielle croit très vite avec T et il est important que L_1 ne soit pas trop grande : il s'agit donc de trouver des fonctions F pour lesquelles l'erreur de consistance $\eta(y, h)$ tende le plus vite possible vers 0 avec h , sans trop détériorer la constante de stabilité du schéma.

DÉMONSTRATION : Comme y est solution de l'EDO $y'(t) = f(t, y(t))$, on a en intégrant cette équation entre les temps nh et $nh + h$

$$y((n+1)h) - y(nh) = \int_{nh}^{nh+h} f(s, y(s)) ds.$$

L'erreur au n-ième pas de temps

$$e_n = y_n - y(nh),$$

vérifie donc par différence

$$e_{n+1} = e_n + h(F(nh, y_n, h) - \frac{1}{h} \int_{nh}^{nh+h} f(s, y(s)) ds)$$

et donc

$$e_{n+1} = e_n + h(\varepsilon_n + \varepsilon'_n)$$

avec

$$\begin{aligned} \varepsilon_n &= F(nh, y_n, h) - F(nh, y(nh), h) \\ \varepsilon'_n &= F(nh, y(nh), h) - \frac{1}{h} \int_{nh}^{nh+h} f(s, y(s)) ds. \end{aligned}$$

Par définition de l'erreur de consistance, on a

$$|\varepsilon'_n| \leq \eta(y, h)$$

et par l'hypothèse de stabilité on a

$$|\varepsilon_n| \leq L_1 |e_n|.$$

On a donc obtenu

$$|e_{n+1}| \leq (1 + L_1 h) |e_n| + h\eta(y, h).$$

Cette inégalité de récurrence se réécrit

$$|e_{n+1}| + r \leq (1 + L_1 h)(|e_n| + r),$$

avec $r = \eta(y, h)/L_1$. On a donc, par récurrence, comme pour le lemme de Gronwall discret 5.2.8 :

$$|e_n| + r \leq (1 + L_1 h)^n (|e_0| + r) \leq \exp(nL_1 h)(|e_0| + r) \leq \exp(TL_1)(|e_0| + r),$$

tant que $nh \leq T$. Comme $e_0 = y_0 - y(0) = a - a = 0$, on a

$$|e_n| \leq (\exp(TL_1) - 1)\eta(y, h)/L_1,$$

ce qui termine la preuve du Théorème 7.1.10. ■

Définition 7.1.12. *S'il existe $k \geq 1$ entier tel que pour f de classe C^k , l'erreur de consistance du schéma est un $O(h^k)$, lorsque h tend vers 0, on dit que le schéma est précis à l'ordre k , ou, plus brièvement, que le schéma est d'ordre k .*

Dans le cas des deux schéma déjà définies, schéma d'Euler et schéma du point milieu, on peut déterminer l'ordre du schéma. On a besoin d'un lemme :

Lemme 7.1.13. *Soit k entier positif ou nul. Si f est de classe C^k , alors toute solution y de l'EDO $y'(t) = f(t, y(t))$ est de classe C^{k+1} .*

DÉMONSTRATION : En effet, comme f est continue et que $y'(t) = f(t, y(t))$, y' est continue. Donc y est de classe C^1 . Si f est de classe C^1 , la fonction composée $g(t) = f(t, y(t))$ est également de classe C^1 . Donc y' est aussi de classe C^1 , ce qui montre qu'en fait y est de classe C^2 . Si f est de classe C^2 , par composition, g se trouve être aussi de classe C^2 . Donc y' est de classe C^2 . Ainsi y est de classe C^3 . On continue ainsi le raisonnement par une récurrence immédiate. ■

Commençons par le schéma d'Euler, 5.1.15 :

Proposition 7.1.14. *On suppose f de classe C^1 vérifiant l'hypothèse (5.2.3). Soit y l'unique solution de l'EDO $y'(t) = f(t, y(t))$ telle que $y(0) = a$ et soit la solution approchée y_n définie par le schéma d'Euler. Alors on a convergence du schéma, avec une erreur d'ordre 1 :*

$$\sup_{nh \leq T} |y(nh) - y_n| \leq M_2 \frac{h}{2}$$

DÉMONSTRATION : Si f est de classe C^1 , alors la solution y est de classe C^2 et on a, par la formule de Taylor à l'ordre 2 :

$$\left| f(t, y(t)) - \frac{y(t+h) - y(t)}{h} \right| = \left| y'(t) - \frac{y(t+h) - y(t)}{h} \right| \leq \frac{h}{2} \sup_{s \in [0, T]} |y''(s)|$$

D'où $\eta(y, h) \leq \frac{h}{2} M_2$ et la méthode est bien d'ordre 1. ■

Pour le schéma du point milieu défini en 7.1.4 l'erreur est d'ordre deux :

Proposition 7.1.15. *On suppose f de classe C^2 vérifiant l'hypothèse (5.2.3). Soit y l'unique solution de l'EDO $y'(t) = f(t, y(t))$ telle que $y(0) = a$ et soit la solution approchée y_n définie par le schéma du point milieu. Alors on a convergence du schéma, avec décroissance quadratique de l'erreur*

$$\sup_{nh \leq T} |y(nh) - y_n| \leq Ch^2$$

où C dépend de L, T et y . Le schéma est d'ordre 2.

DÉMONSTRATION : Si y est la solution de l'équation $y'(t) = f(t, y(t))$ telle que $y(0) = a$, on a, pour le schéma du point milieu :

$$\begin{aligned} F(t, y(t), h) - \frac{y(t+h) - y(t)}{h} &= f(t + h/2, y(t) + h/2f(t, y(t))) - \frac{y(t+h) - y(t)}{h} \\ &= f(t + h/2, y(t) + h/2y'(t)) - \frac{y(t+h) - y(t)}{h} = r + \tilde{r}, \end{aligned}$$

où

$$\begin{aligned} r &= f(t + h/2, y(t) + h/2y'(t)) - f(t + h/2, y(t + h/2)) \\ \tilde{r} &= f(t + h/2, y(t + h/2)) - \frac{y(t+h) - y(t)}{h} = y'(t + h/2) - \frac{y(t+h) - y(t)}{h} \end{aligned}$$

En utilisant un développement de Taylor autour du point $t + h/2$, on trouve, comme au chapitre 3 :

$$y(t+h) - y(t + \frac{h}{2}) = \frac{h}{2}f(t + h/2, y(t + h/2)) + \frac{h^2}{8}y''(t + \frac{h}{2}) + \frac{h^3}{48}y'''(s_1), \quad t + \frac{h}{2} \leq s_1 \leq t + h$$

$$y(t) - y(t + \frac{h}{2}) = -\frac{h}{2}f(t + h/2, y(t + h/2)) + \frac{h^2}{8}y''(t + \frac{h}{2}) - \frac{h^3}{48}y'''(s_2), \quad t \leq s_2 \leq t + \frac{h}{2}$$

D'où :

$$\tilde{r} = \frac{h^3}{48}y'''(s_1) - \frac{h^3}{48}y'''(s_2)$$

et par suite,

$$|\tilde{r}| \leq \frac{h^2}{24} \sup_{s \in [0, T]} |y'''(s)|$$

En utilisant la condition de Lipschitz (5.2.3), on obtient

$$|r| \leq L|y(t) + h/2y'(t) - y(t + h/2)|$$

et donc, encore par développement de Taylor mais autour du point t :

$$|r| \leq L \frac{h^2}{8} \sup_{s \in [0, T]} |y''(s)|.$$

Ainsi on a obtenu

$$\eta(y, h) = \sup_{0 \leq t \leq t+h \leq T} |F(t, y(t), h) - \frac{1}{h} \int_t^{t+h} f(s, y(s)) ds| \leq Ch^2,$$

avec C donné par $C = \frac{1}{24} (\sup_{0 \leq t \leq T} |y'''(t)| + 3L \sup_{0 \leq t \leq T} |y''(t)|)$. Ceci prouve la Proposition en utilisant le Théorème 7.1.10 ■

Si la fonction F est assez régulière, on peut donner une condition nécessaire et suffisante pour qu'un schéma numérique explicite soit d'ordre supérieur ou égal à k :

Proposition 7.1.16. Soit $y'(t) = f(t, y(t))$ une EDO et y est la solution de l'EDO telle que $y(0) = y_0$ et soit $F : (t, x, h) \in [0, T] \times \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}^m$, de classe C^k , vérifiant l'hypothèse de consistance. Le schéma explicite $y_{n+1} = y_n + hF(nh, y_n, h)$ est d'ordre $\geq k$ si et seulement si, pour tout $l \leq k - 1$:

$$\frac{\partial^l F}{\partial h^l}(t, y(t), 0) = \frac{1}{l+1} f^{[l]}(t, y(t))$$

où on définit par récurrence $f^{[l]}(t, x)$ pour $x \in \mathbb{R}^m$:

$$f^{[0]}(t, x) = f(t, x), \dots, f^{[l+1]}(t, x) = \frac{\partial f^{[l]}}{\partial t}(t, x) + \frac{\partial f^{[l]}}{\partial x}(t, x)f(t, x), \quad 0 \leq l \leq k - 1.$$

Cette condition est en particulier réalisée si $\forall x \in \mathbb{R}^m$,

$$\frac{\partial^l F}{\partial h^l}(t, x, 0) = \frac{1}{l+1} f^{[l]}(t, x)$$

DÉMONSTRATION : La formule de Taylor, appliquée à la fonction $h \rightarrow F(t, x, h)$ en $h = 0$ à l'ordre k donne :

$$F(t, x, h) = \sum_{l=0}^{k-1} \frac{1}{l!} h^l \frac{\partial^l F}{\partial h^l}(t, x, 0) + O(h^k)$$

D'autre part, puisque la fonction f est de classe C^k , la solution y de l'EDO $y'(t) = f(t, y(t))$ est de classe C^{k+1} et on vérifie aisément par récurrence que

$$f^{[l]}(t, y(t)) = \frac{d}{dt} f^{[l-1]}(t, y(t)) = y^{(l+1)}(t)$$

La formule de Taylor appliquée à la fonction y en t à l'ordre $k + 1$ donne :

$$y(t+h) - y(t) = \sum_{p=1}^k \frac{1}{p!} h^p y^{(p)}(t) + O(h^{k+1}) = \sum_{l=0}^{k-1} \frac{1}{l+1!} h^{l+1} f^{[l]}(t, y(t)) + O(h^{k+1})$$

D'où

$$\eta(y, h) = \sup_{0 \leq t \leq T-h} \left| \sum_{l=0}^{k-1} \frac{1}{l!} h^l \left[\frac{1}{l+1} f^{[l]}(t, y(t)) - \frac{\partial^l F}{\partial h^l}(t, y(t), 0) \right] \right| + O(h^k)$$

ce qui prouve la Proposition 7.1.16. ■

Donnons maintenant quelques exemples de schémas explicites :

Exemples 7.1.17. Cas d'une équation différentielle linéaire autonome $y'(t) = Ay(t)$ avec condition initiale $y(0) = a$:

- 1) Le schéma d'Euler est donné par $y_{n+1} = (I + hA)y_n$, $y_0 = a$
- 2) Le schéma du point milieu est donné par $y_{n+1} = (I(1 + h) + \frac{h^2}{2}A)y_n$, $y_0 = a$
- 3) On peut aussi définir le schéma exact, tel que $F(t, x, h) = \frac{1}{h}(e^{hA}x - x)$ ce qui donne : $y_{n+1} = e^{hA}y_n$, $y_0 = a$.

Pour compléter ces exemples, on va donner un schéma d'approximation dans le cas particulier où la fonction f est la somme de deux fonctions Φ et Ψ :

Exemple 7.1.18. Schéma de décomposition : Supposons que l'on ait deux EDO, respectivement

$$y'(t) = \phi(t, y(t)),$$

$$y'(t) = \psi(t, y(t)),$$

et deux schémas numériques explicites associés définis respectivement par

$$y_{n+1} = y_n + h\Phi(nh, y_n, h),$$

$$y_{n+1} = y_n + h\Psi(nh, y_n, h),$$

où les fonctions Φ et Ψ supposées définies et continues sur $[0, T] \times \mathbb{R}^m \times [0, T]$ vérifient les hypothèses de consistance et stabilité

$$\Phi(t, x, 0) = \phi(t, x), \quad |\Phi(t, b, h) - \Phi(t, c, h)| \leq L_1|b - c|,$$

$$\Psi(t, x, 0) = \psi(t, x), \quad |\Psi(t, b, h) - \Psi(t, c, h)| \leq L_2|b - c|.$$

A partir de cela on peut former un schéma pour l'équation

$$y'(t) = \phi(t, y(t)) + \psi(t, y(t))$$

en posant

$$y_{n+1} = y_* + h\Psi(nh, y_*, h),$$

$$y_* = y_n + h\Phi(nh, y_n, h).$$

Autrement dit, on effectue un pas de la première équation, suivi d'un pas de la deuxième équation.

Il est très facile de vérifier que le schéma de décomposition est stable et consistant, après l'avoir réécrit sous la forme

$$y_{n+1} = y_n + hF(nh, y_n, h)$$

avec

$$F(t, x, h) = \Phi(t, x, h) + \Psi(t, x + h\Phi(t, x, h), h).$$

En revanche, ce schéma ne peut en général dépasser l'ordre 1 en précision.

Il existe une variante, appelée schéma de Strang, permettant de récupérer l'ordre 2 dans la mesure où les deux schémas de départ sont au moins d'ordre 2. Il suffit en effet d'effectuer un demi-pas $h/2$ de la première équation, puis un pas h de la deuxième, suivi de nouveau d'un demi-pas $h/2$ de la première, pour rétablir la précision d'ordre 2.

Autrement dit on pose

$$\begin{aligned} y_{n+1} &= y_{**} + \frac{h}{2}\Phi(nh, y_{**}, h), \\ y_{**} &= y_* + h\Psi(nh, y_*, h). \\ y_* &= y_n + \frac{h}{2}\Phi(nh, y_n, h). \end{aligned}$$

Nous allons maintenant examiner deux applications du schéma de décomposition : la formule de Trotter pour les exponentielles de matrice et l'analyse des systèmes raides.

Application 7.1.19. *La formule de Trotter.*

Un cas très particulier du schéma de décomposition est celui où on sait résoudre exactement chacune des deux équations $y'(t) = \phi(t, y(t))$ et $y'(t) = \psi(t, y(t))$. Prenons par exemple le cas d'équations linéaires autonomes

$$\phi(t, x) = A.x, \quad \psi(t, x) = B.x,$$

où A et B sont des matrices réelles $m \times m$. On peut alors considérer le schéma de décomposition où chacune des deux équations est successivement résolues par le schéma exact, 7.1.17, sur un pas de longueur h , ce qui donne, avec la notation d'exponentielles de matrices déjà vue,

$$y_{n+1} = \exp(hB).y_* = \exp(hB).\exp(hA).y_n$$

et correspond aux choix suivants de Φ et Ψ

$$\Phi(t, x, h) = \frac{1}{h}(\exp(hA).x - x), \quad \Psi(t, x, h) = \frac{1}{h}(\exp(hB).x - x),$$

pour $h > 0$ avec

$$\Phi(t, x, 0) = A.x, \quad \Psi(t, x, 0) = B.x.$$

Les conditions de stabilité et de consistance sont satisfaites. Pour la stabilité, on pourra utiliser la majoration suivante :

$$\begin{aligned} \left| \frac{1}{h}(\exp(hA).x - x) \right| &= \left| \sum_{k=1}^{\infty} \frac{h^{k-1}A^k}{k!}x \right| \\ &\leq \sum_{k=1}^{\infty} \frac{h^{k-1}\|A\|^k}{k!}|x| \\ &\leq \sum_{k=1}^{\infty} \frac{h^{k-1}\|A\|^k}{(k-1)!}|x| = \|A\| \exp(h\|A\|)|x| \end{aligned}$$

Par récurrence, on obtient :

$$y_n = (\exp(hB) \cdot \exp(hA))^n \cdot a,$$

où a est la donnée initiale du problème.

Du théorème général de convergence des schémas explicites, on déduit, en posant $h = T/n$ avec $n > 0$ entier que l'unique solution y de

$$y'(t) = (A + B) \cdot y(t), \quad y(0) = a$$

est égale, au temps $t = T$, à la limite de cette équation lorsque $n \rightarrow \infty$.

Or cette solution est égale à $\exp((A+B)T) \cdot a$, par définition de l'exponentielle d'une matrice. Comme on peut choisir a de façon arbitraire et qu'on peut poser $T = 1$, on a ainsi obtenu le résultat suivant sur les exponentielles de matrice :

Corollaire 7.1.20. (Formule de Trotter) Soit A et B deux matrices réelles $m \times m$. Alors, pour tout $a \in \mathbb{R}^m$,

$$\exp(A + B) \cdot a = \lim_{n \rightarrow \infty} (\exp(B/n) \cdot \exp(A/n))^n \cdot a.$$

Cette formule est très utile en pratique. Elle est bien entendu triviale dans le cas très spécial où $m = 1$ et donc que A et B sont des scalaires, à cause des propriétés de la fonction exponentielle. Elle ne l'est plus du tout dès que $m \geq 2$, sauf dans le cas particulier où les matrices A et B commutent ($AB = BA$). Dans ce dernier cas, on peut directement écrire la formule

$$\exp(A + B) = \exp(B) \cdot \exp(A) = \exp(A) \cdot \exp(B)$$

qui est fautive lorsque A et B ne commutent pas.

Cas des Systèmes raides. On considère un système de la forme

$$y'(t) = \beta A y(t) + g(t, y(t)),$$

où $\beta > 0$ est un réel, A est une matrice $m \times m$ réelle fixée et la fonction g est définie et continue pour tous $t \geq 0$, satisfait la condition de Lipschitz 5.2.3, uniformément en $t \geq 0$. On suppose de plus que g est bornée

$$|g(t, x)| \leq M, \quad \forall t \geq 0, \forall x \in \mathbb{R}^m.$$

On suppose aussi que A est soit une matrice antisymétrique (**cas 1**), soit une matrice symétrique à valeurs propres négatives ou nulles (**cas 2**). Lorsque β est grand par rapport à M , on dit que le système est raide. C'est une situation très fréquente dans les applications. Comme exemple, très important, de systèmes raides, citons ceux qu'on obtient à partir de la discrétisation en espace par différences finies d'équations aux dérivées partielles. Prenons l'exemple de l'équation dite de réaction-diffusion qui modélise de nombreux phénomènes en biologie, chimie, finance etc...

$$\frac{\partial}{\partial t} u(t, x) = \frac{\partial^2}{\partial x^2} u(t, x) + g(t, x),$$

où $g(t, x)$ est une fonction donnée et l'inconnue est une fonction $u(t, x)$ à déterminer pour $t \in [0, T]$ et $x \in [0, L]$, en fixant sa valeur en $t = 0$ $u(0, x) = u_0(x)$, en $x = 0$ et $x = L$, par exemple $u(t, 0) = u(t, L) = 0$.

L'idée est de résoudre approximativement cette équation aux points $x_i = iL/(N + 1)$ pour $i = 1, \dots, N$ où N est un entier grand fixé et d'approcher la solution exacte, dont on peut prouver qu'elle est bien définie, $u(t, x_i)$ par $y_i(t)$ définie comme suit. On approche la dérivée partielle $\frac{\partial^2}{\partial x^2} u(t, x_i)$ au temps t et au point x_i par

$$(y_{i+1}(t) - 2y_i(t) + y_{i-1}(t))\beta$$

avec $\beta = (N + 1)^2/L^2$. On demande alors au vecteur $Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_N(t) \end{pmatrix}$ de vérifier l'EDO

$\frac{dY}{dt} = \beta AY + g(t, Y)$, où A est la matrice $N \times N$ définie par

$$A = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 & 0 \\ \dots & & & & & & & \\ \dots & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & \dots & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & -2 \end{pmatrix}$$

dont on peut montrer qu'elle est symétrique à valeurs propres strictement négatives.

Enfin, la condition initiale est donnée par : $y(0) = (u_0(x_1), \dots, u_0(x_N))$.

Après cet exemple, examinons, dans le cadre général des systèmes raides, le schéma de décomposition suivant, exact pour la partie linéaire, basé sur le schéma d'Euler pour le reste

$$y_{n+1} = \exp(h\beta A)y_*$$

$$y_* = y_n + hg(nh, y_n).$$

Proposition 7.1.21. 1) Sous les hypothèses précédentes, et dans les deux cas 1 et 2, on a pour tout $s \geq 0$,

$$\|\exp(sA)\| \leq 1,$$

où $\|\cdot\|$ est la norme matricielle associée à la norme euclidienne sur \mathbb{R}^m .

2) La solution exacte y de $y'(t) = \beta Ay(t) + g(t, y(t))$, avec pour donnée initiale $y_0 \in \mathbb{R}^m$, vérifie

$$|y(t)| \leq |y_0| + tM, \quad \forall t \geq 0$$

et la solution approchée donnée par le schéma de décomposition des systèmes raides

$$|y_n| \leq |y_0| + nhM, \quad \forall n = 0, 1, 2, \dots$$

DÉMONSTRATION : Montrer 1) revient exactement à montrer 2) dans le cas particulier $g = 0$ et $M = 0$. En effet dans ce cas, la solution est donnée par $y(t) = \exp(t\beta A).y_0$. Or dans ce cas, on a

$$\frac{d}{dt}|y(t)|^2 = \langle 2y(t), y'(t) \rangle = \langle 2y(t), \beta Ay(t) \rangle$$

Le terme de droite est nul lorsque A est antisymétrique et négatif ou nul si A est symétrique semi-définie négative. Dans les deux cas, $|y(t)|^2$ est donc une fonction décroissante de t , ce qui montre bien 2) avec $M = 0$.

Pour montrer 2) dans le cas général, il suffit de prouver l'inégalité discrète $|y_n| \leq |y_0| + nhM$, $\forall n = 0, 1, 2, \dots$, puisque la théorie générale nous prouve la convergence du schéma de décomposition, ce qui permet d'obtenir l'inégalité exacte $|y(t)| \leq |y_0| + tM$, $\forall t \geq 0$ en passant à la limite $h \rightarrow 0$ dans l'inégalité discrète. Or, l'inégalité discrète s'obtient facilement à partir de l'expression $y_{n+1} = \exp(h\beta A)y_*$, $y_* = y_n + hg(nh, y_n)$, en utilisant la bornitude de g et 1).

En effet, on a

$$|y_{n+1}| = |\exp(h\beta A).y_*| \leq |y_*| = |y_n + hg(nh, y_n)| \leq |y_n| + hM$$

et le résultat en découle par récurrence sur n . ■

Remarque. Il est remarquable que les bornes ne dépendent pas de la taille du paramètre β dans 2). Ceci ne serait plus vrai en général si on se contentait d'utiliser le schéma d'Euler pour la partie linéaire du schéma de décomposition, ce qui donnerait

$$\begin{aligned} y_{n+1} &= (1 + h\beta A)y_* \\ y_* &= y_n + hg(nh, y_n). \end{aligned}$$

7.2 Equations de type gradient et schémas implicites

Dans cette section, on va étudier un exemple de schéma implicite. Comme pour les schémas explicites, on se donne une fonction F

$$(t, x, h) \in [0, T] \times \mathbb{R}^m \times [0, T] \rightarrow F(t, x, h) \in \mathbb{R}^m$$

qu'on suppose continue et qui vérifie les deux propriétés de consistance et de stabilité.

On définit alors le schéma implicite par :

Définition 7.2.1. Soit $0 < h \leq T$ un pas de temps. Le schéma implicite est donné par la récurrence suivante :

$$y_{n+1} = y_n + hF(nh, y_{n+1}, h), \quad n = 0, 1, 2, \dots, \quad y_0 = a,$$

où a est la donnée initiale et tant que $nh \leq T$.

Un tel schéma est dit implicite car il ne donne pas une valeur explicite de y_{n+1} en fonction de y_n et nécessite la résolution d'une équation. Il n'est d'ailleurs pas forcément bien défini.

L'existence d'une fonction de Liapounov généralisée est un bon cadre pour l'étude théorique des schémas implicites dont le plus simple est le schéma dit d'Euler implicite :

Définition 7.2.2. Etant donnée une EDO $y'(t) = f(t, y(t))$, le schéma dit d'Euler implicite est donné par

$$y_{n+1} = y_n + hf(t, y_{n+1}), \quad n = 0, 1, 2, \dots,$$

où y_0 est la donnée initiale et $h > 0$ le pas de discrétisation en temps.

Exemple 7.2.3. Cas d'une équation différentielle linéaire autonome $y'(t) = Ay(t)$.

Le schéma d'Euler implicite est défini par

$$y_{n+1} = (1 - h\beta A)^{-1}y_n$$

Dans ce cas particulier, on voit que l'on a à inverser la matrice $I - hA$ ce qui n'est pas toujours possible : par exemple, si A a une valeur propre réelle positive, le schéma n'est plus bien défini si on prend pour h cette valeur propre.

C'est néanmoins toujours possible dans les cas très fréquemment rencontrés où A est soit symétrique négative soit antisymétrique.

Cela dit, en pratique, l'inversion numérique nécessite en général beaucoup d'effort ou une bonne pratique des bibliothèques de logiciels de calcul scientifique.

Lorsque $V = \mathbb{R}^m$, on peut montrer le résultat théorique suivant :

Proposition 7.2.4. Soit une EDO autonome de type gradient $y'(t) = f(y(t))$ avec $f = -\nabla U$ où U est une fonction convexe deux fois continument différentiable sur \mathbb{R}^m tendant vers l'infini à l'infini (i.e. telle que $U(x) \rightarrow +\infty$ lorsque $|x| \rightarrow +\infty$). Alors U est une fonction de Liapounov, au sens strict de (5.3.8 2)), le schéma d'Euler implicite est bien défini et la solution approchée vérifie l'estimation

$$U(y_{n+1}) \leq U(y_n) \leq \dots U(y_0), \quad n = 1, 2, \dots,$$

quel que soit le choix du pas $h > 0$.

Cette estimation est parfois appelée propriété de stabilité au sens de Liapounov, inconditionnelle car indépendante de la taille de h . Le schéma d'Euler explicite n'a pas cette propriété : on peut le vérifier déjà sur le cas le plus simple où $m = 1$, $f(x) = -x$ et $U(x) = x^2/2$.

Plus généralement, ces hypothèses sont vérifiées dans le cas où U est une fonction quadratique de la forme $\langle Ax, x \rangle$ avec A matrice symétrique définie positive.

DÉMONSTRATION : Notons d'abord que dire que U , qui est continue, tend vers l'infini à l'infini entraîne que les ensembles de niveau 5.3.8 1) sont des parties compactes de \mathbb{R}^m . Ainsi U est bien fonction de Liapounov.

Pour montrer que le schéma est bien défini, il suffit de montrer que l'équation

$$x - hf(x) = z,$$

a une unique solution $x \in \mathbb{R}^m$ pour tout $h > 0$ et $z \in \mathbb{R}^m$ fixés. Or cette équation est équivalente à

$$x + h\nabla U(x) = z, \quad \text{ou encore} \quad D_x\left(\frac{1}{2}|x - z|^2 + hU(x)\right) = 0,$$

où on note D_x la différentielle par rapport à x .

Ceci équivaut à rechercher l'unique minimum de la fonction convexe qui tend vers l'infini à l'infini :

$$x \rightarrow \frac{1}{2}|x - z|^2 + hU(x).$$

Pour montrer l'estimation de 7.2.4, on utilise l'une des caractérisations possible des fonctions convexes continument différentiables, à savoir

$$U(b) \geq U(c) + DU(c).(b - c), \quad \forall b, c \in \mathbb{R}^m,$$

qui dit que le graphe d'une fonction convexe est toujours au dessus de ses hyperplans tangents. De cette propriété, on déduit

$$U(y_n) \geq U(y_{n+1}) + DU(y_{n+1}).(y_n - y_{n+1}),$$

d'où, en utilisant que $f = -\nabla u$ et la définition du schéma d'Euler implicite

$$U(y_n) \geq U(y_{n+1}) + h|\nabla U(y_{n+1})|^2.$$

Cette inégalité implique le résultat quel que soit le choix de $h > 0$. ■

7.3 Méthodes de Runge-Kutta

Ces méthodes sont des schémas implicites particuliers, où le choix de la fonction F est déterminé à partir de méthodes de quadrature.

On considère l'EDO avec condition initiale, définie sur $t \in [t_0, t_0 + T]$ par :

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

On se donne $q \in \mathbb{N}$ et pour $1 \leq i \leq q$, des coefficients $c_i \in [0, 1]$ tels que $0 = c_1 \leq c_2 \leq \dots \leq c_q \leq 1$.

On considère une subdivision $t_0 < t_1 < \dots < t_N = t_0 + T$ de l'intervalle $[0, T]$, en posant pour chaque $0 \leq n \leq N$, $h_n = t_{n+1} - t_n$ et en associant à chacun de ces points une valeur y_n et une pente $p_n = f(t_n, y_n)$.

Pour chaque $0 \leq n \leq N$, on considère des points intermédiaires $t_{n,i}$, définis par

$$t_{n,i} = t_n + c_i h_n, \quad 1 \leq i \leq q.$$

A chacun de ces points, on associe une valeur $y_{n,i}$ et la pente correspondante définie par $p_{n,i} = f(t_{n,i}, y_{n,i})$.

Comme $c_1 = 0$, on aura en particulier : $t_{n,1} = t_n$, $y_{n,1} = y_n$, $p_{n,1} = p_n = f(t_n, y_n)$.

Soit y une solution exacte de l'équation EDO . On peut écrire, pour $0 \leq n \leq N - 1$ et $1 \leq i \leq q$:

$$y(t_{n,i}) = y(t_n) + \int_{t_n}^{t_{n,i}} f(t, y(t)) dt = y(t_n) + h_n \int_0^{c_i} f(t_n + uh_n, y(t_n + uh_n)) du$$

et de même :

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt = y(t_n) + h_n \int_0^1 f(t_n + uh_n, y(t_n + uh_n)) du$$

grâce au changement de variable $t = t_n + uh_n$.

On se donne alors pour chaque $i = 2, \dots, q$ des méthodes de quadrature (M_i) sur $[0, c_i]$ et une méthode de quadrature (M) sur $[0, 1]$, ces méthodes pouvant être a priori différentes :

$$(M_i) \int_0^{c_i} g(u) du \sim \sum_{1 \leq j < i} a_{i,j} g(c_j), \text{ avec } \sum_{1 \leq j < i} a_{i,j} = c_i$$

$$(M) \int_0^1 g(u) du \sim \sum_{1 \leq i \leq q} b_i g(c_i), \text{ avec } \sum_{1 \leq i \leq q} b_i = 1$$

En appliquant ces méthodes à la fonction $g(u) = f(t_n + uh_n, y(t_n + uh_n))$, on trouve alors :

$$y(t_{n,i}) \sim y(t_n) + h_n \sum_{1 \leq j < i} a_{i,j} f(t_{n,j}, y(t_{n,j}))$$

$$y(t_{n+1}) \sim y(t_n) + h_n \sum_{1 \leq i \leq q} b_i f(t_{n,i}, y(t_{n,i}))$$

A partir de ces approximations, on peut définir la méthode de Rung-Kutta correspondante par l'algorithme :

$$\left\{ \begin{array}{l} t_{n,i} = t_n + c_i h_n \\ y_{n,i} = y_n + h_n \sum_{1 \leq j < i} a_{i,j} p_{n,j} \\ p_{n,i} = f(t_{n,i}, y_{n,i}) \end{array} \right\} 1 \leq i \leq q$$

$$\left\{ \begin{array}{l} t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n \sum_{1 \leq i \leq q} b_i p_{n,i} \end{array} \right\}$$

On représente cette méthode par le tableau de nombres suivant, à q lignes et $q + 1$ colonnes :

$$\begin{array}{ccccccc} (M_2) & c_2 & a_{2,1} & 0 & \dots & 0 & 0 \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ (M_q) & c_q & a_{q,1} & a_{q,2} & \dots & a_{q,q-1} & 0 \\ (M) & 1 & b_1 & b_2 & \dots & b_{q-1} & b_q \end{array}$$

où les méthodes d'intégration approchées correspondent aux lignes.

Les méthodes de Runge-Kutta sont des méthodes à un pas, c'est à dire vérifiant un schéma explicite

$$y_{n+1} = y_n + h_n F(t_n, y_n, h_n)$$

avec $F(t_n, y_n, h_n) = \sum_{1 \leq i \leq q} b_i p_{n,i}$, c'est à dire de manière plus explicite, pour $((t, x, h) \in [0, T] \times \mathbb{R} \times [0, T])$:

$$(*) \quad \begin{cases} F(t, x, h) = \sum_{1 \leq i \leq q} b_i f(t + c_i h, x_i) \\ x_i = x + h \sum_{1 \leq j < i} a_{i,j} f(t + c_j h, x_j), \quad 1 \leq i \leq q \end{cases}$$

Exemple 7.3.1. $q = 1$.

Le seul tableau possible est

$$(M) \quad \begin{matrix} & 1 & 1 \end{matrix}$$

c'est à dire $b_1 = 1$. L'algorithme correspondant est donné par :

$$\begin{cases} p_{n,1} = p_n = f(t_n, y_n) \\ t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n p_n \end{cases}$$

Il s'agit de la méthode d'Euler.

Exemple 7.3.2. $q = 2$.

Pour $\alpha \in]0, 1]$, on considère les tableaux de la forme

$$(M_2) \quad \begin{matrix} & \alpha & \alpha & 0 \\ (M) & 1 & 1 - \frac{1}{2\alpha} & \frac{1}{2\alpha} \end{matrix}$$

L'algorithme correspondant s'écrit :

$$\begin{cases} p_{n,1} = p_n = f(t_n, y_n) \\ t_{n,2} = t_n + \alpha h_n \\ y_{n,2} = y_n + \alpha h_n p_{n,1} \\ p_{n,2} = f(t_{n,2}, y_{n,2}) \\ t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n [(1 - \frac{1}{2\alpha})p_{n,1} + \frac{1}{2\alpha}p_{n,2}] \end{cases}$$

Sous forme condensée, ceci donne :

$$y_{n+1} = y_n + h_n [(1 - \frac{1}{2\alpha})f(t_n, y_n) + \frac{1}{2\alpha}f(t_n + \alpha h_n, y_n + \alpha h_n f(t_n, y_n))]$$

- Pour $\alpha = \frac{1}{2}$, on retrouve la méthode du point milieu :

$$y_{n+1} = y_n + h_n [f(t_n + \frac{h_n}{2}, y_n + \frac{h_n}{2} f(t_n, y_n))]$$

qui repose sur la méthode d'intégration du point milieu, 3.1.4 :

$$\int_0^1 g(u) du \sim g\left(\frac{1}{2}\right)$$

• Pour $\alpha = 1$, on obtient la méthode de Heun :

$$y_{n+1} = y_n + h_n \left[\frac{1}{2} f(t_n, y_n) + \frac{1}{2} f(t_{n+1}, y_n + \alpha h_n f(t_n, y_n)) \right]$$

qui repose sur la méthode des trapèzes 3.2.1

$$\int_0^1 g(u) du \sim \frac{1}{2} [g(0) + g(1)]$$

Exemple 7.3.3. Méthode de Runge-Kutta classique, $q = 4$.

On considère le tableau suivant :

(M_2)	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
(M_3)	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
(M_4)	1	0	0	1	0
(M)	1	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

L'algorithme correspondant s'écrit :

$$\left\{ \begin{array}{l} p_{n,1} = p_n = f(t_n, y_n) \\ t_{n,2} = t_n + \frac{1}{2}h_n \\ y_{n,2} = y_n + \frac{1}{2}h_n p_{n,1} \\ p_{n,2} = f(t_{n,2}, y_{n,2}) \\ y_{n,3} = y_n + \frac{1}{2}h_n p_{n,2} \\ p_{n,3} = f(t_{n,2}, y_{n,3}) \\ y_{n,4} = y_n + h_n p_{n,3} \\ p_{n,4} = f(t_{n+1}, y_{n,4}) \\ t_{n+1} = t_n + h_n \\ y_{n+1} = y_n + h_n \left[\frac{1}{6}p_{n,1} + \frac{2}{6}p_{n,2} + \frac{2}{6}p_{n,2} + \frac{1}{6}p_{n,4} \right] \end{array} \right.$$

On notera qu'ici, $t_{n,3} = t_{n,2}$ et $t_{n,4} = t_{n+1}$.

Dans ce cas, les méthodes d'intégration utilisées sont respectivement :

$$(M_2) \quad \int_0^{\frac{1}{2}} g(u) du \sim \frac{1}{2} g(0) \quad \text{rectangles à gauche}$$

$$(M_3) \quad \int_0^{\frac{1}{2}} g(u) du \sim \frac{1}{2} g\left(\frac{1}{2}\right) \quad \text{rectangles à droite}$$

$$(M_4) \quad \int_0^1 g(u) du \sim g\left(\frac{1}{2}\right) \quad \text{point milieu}$$

$$(M) \quad \int_0^1 g(u) du \sim \frac{1}{6}g(0) + \frac{2}{6}g\left(\frac{1}{2}\right) + \frac{2}{6}g\left(\frac{1}{2}\right) + \frac{1}{6}g(1) \quad \text{Simpson}$$

Pour que les méthodes de Rung Kutta rentrent dans le cadre général des schémas explicites à un pas, il faut que la fonction F , définie par la formule (*) vérifie les hypothèses de consistance et de stabilité, 7.1.1. Il est clair que l'hypothèse de consistance est vérifiée par construction. Pour l'hypothèse de stabilité, on a le résultat suivant :

Proposition 7.3.4. *Si f est k lipschitzienne, la fonction F définie par la formule (*) est aussi lipschitzienne. Précisément, soit $\alpha = \max\left\{\sum_{1 \leq j < i} |a_{i,j}|, 1 \leq i \leq q\right\}$, alors si z et y sont deux réels :*

$$|F(t, y, h) - F(t, z, h)| \leq \sum_{1 \leq i \leq q} |b_i| k |y_i - z_i| \leq \Lambda |y - z|$$

avec $\Lambda = k \sum_{1 \leq i \leq q} |b_i| [1 + \alpha hk + \dots + (\alpha hk)^{i-1}]$.

DÉMONSTRATION : On a besoin d'un lemme :

Lemme 7.3.5. *Avec les notations du théorème 7.3.4,*

$$|y_i - z_i| \leq [1 + \alpha hk + \dots + (\alpha hk)^{i-1}] |y - z|.$$

On démontre ce lemme par récurrence sur i : pour $i = 1$, on a $y_1 = y$ et $z_1 = z$, le résultat est évident. Supposons l'égalité vraie pour tout $j < i$. Alors :

$$|y_i - z_i| \leq |y - z| + h \sum_{j < i} |a_{i,j}| k \max_{j < i} |y_j - z_j|$$

$$|y_i - z_i| \leq |y - z| + \alpha hk \max_{j < i} |y_j - z_j|$$

En appliquant l'hypothèse de récurrence, on a

$$\max_{j < i} |y_j - z_j| \leq (1 + \alpha hk + \dots + (\alpha hk)^{i-2}) |y - z|$$

et l'inégalité à l'ordre i s'en déduit immédiatement.

La formule (*) entraîne alors :

$$|F(t, y, h) - F(t, z, h)| \leq \sum_{1 \leq i \leq q} |b_i| k |y_i - z_i| \leq \Lambda |y - z|$$

■

On va s'intéresser à l'ordre de ces méthodes, en utilisant la Proposition 7.1.16 :

Théorème 7.3.6. Si la fonction f est de classe C^3 , les méthodes de Rung-Kutta définies par le tableau des coefficients c_i, a_{ij}, b_j sont :

$$\text{d'ordre } \geq 1 \text{ dès que } \sum_{1 \leq i \leq q} b_i c_i = \frac{1}{2}$$

$$\text{d'ordre } \geq 2 \text{ dès que } \sum_{1 \leq i \leq q} b_i c_i = \frac{1}{2}, \quad \sum_{1 \leq i \leq q} b_i c_i^2 = \frac{1}{3}, \quad \sum_{1 \leq i \leq q, 1 \leq j < i} b_i a_{ij} c_j = \frac{1}{6}$$

$$\text{d'ordre } \geq 3 \text{ dès que } \sum_{1 \leq i \leq q} b_i c_i = \frac{1}{2}, \quad \sum_{1 \leq i \leq q} b_i c_i^2 = \frac{1}{3}, \quad \sum_{1 \leq i \leq q} b_i c_i^3 = \frac{1}{4},$$

$$\sum_{1 \leq i \leq q, 1 \leq j < i} b_i a_{ij} c_j = \frac{1}{6}, \quad \sum_{1 \leq i \leq q, 1 \leq j < i} b_i a_{ij} c_j^2 = \frac{1}{12}, \quad \sum_{1 \leq i \leq q, 1 \leq j < i} b_i c_i a_{ij} c_j = \frac{1}{8}$$

$$\sum_{1 \leq i \leq q, 1 \leq j < i, 1 \leq k < i} b_i a_{ij} a_{jk} c_k = \frac{1}{12}.$$

DÉMONSTRATION : Les calculs sont fastidieux, nous ne les effectuerons que pour les ordres 1 et 2. Il s'agit de calculer les dérivées partielles $\frac{\partial^l F}{\partial h^l}$ de la fonction $h \rightarrow F(t, x, h)$ définie par (*), pour $l = 1, 2, 3$.

On rappelle que la fonction F est définie par :

$$(*) \quad \begin{cases} F(t, x, h) = \sum_{1 \leq i \leq q} b_i f(t + c_i h, x_i) \\ x_i = x + h \sum_{1 \leq j < i} a_{i,j} f(t + c_j h, x_j), \quad 1 \leq i \leq q \end{cases}$$

A l'ordre 0, on a $F(t, x, 0) = f(t, x)$, ce qui traduit simplement la consistance de F .

A l'ordre 1, on a :

$$\frac{\partial F}{\partial h}(t, x, h) = \sum_{1 \leq i \leq q} b_i \left(c_i \frac{\partial f}{\partial t}(t + c_i h, x_i) + \frac{\partial f}{\partial x}(t + c_i h, x_i) \frac{\partial x_i}{\partial h} \right)$$

$$\frac{\partial x_i}{\partial h} = \sum_{1 \leq j < i} a_{i,j} f(t + c_j h, x_j) + h \sum_{1 \leq j < i} a_{i,j} \left(\frac{\partial f}{\partial t}(t + c_j h, x_j) + \frac{\partial f}{\partial x}(t + c_j h, x_j) \frac{\partial x_j}{\partial h} \right)$$

Donc pour $h = 0$: $x_i = x$ et $\frac{\partial x_i}{\partial h}|_{h=0} = \sum_{1 \leq j < i} a_{i,j} f(t, x) = c_i f(t, x)$, ce qui donne :

$$\begin{aligned} \frac{\partial F}{\partial h}(t, x, 0) &= \sum_{1 \leq i \leq q} b_i \left(c_i \frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x) \frac{\partial x_i}{\partial h}|_{h=0} \right) \\ &= \sum_{1 \leq i \leq q} b_i c_i \left(\frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x) f(t, x) \right) = \sum_{1 \leq i \leq q} b_i c_i f^{[1]}(t, x) \end{aligned}$$

D'après la Proposition 7.1.16, la méthode est d'ordre ≥ 1 dès que $\sum_{1 \leq i \leq q} b_i c_i = \frac{1}{2}$.

A l'ordre 2, on calcule la dérivée seconde de F par rapport à h selon le même principe, ce qui donne en abrégant un peu les calculs :

$$\begin{aligned} \frac{\partial^2 F}{\partial h^2}(t, x, h) &= \sum_{1 \leq i \leq q} b_i (c_i^2 \frac{\partial^2 f}{\partial t^2}(t + c_i h, x_i) + 2c_i \frac{\partial^2 f}{\partial t \partial x}(t + c_i h, x_i) \frac{\partial x_i}{\partial h} \\ &\quad + \frac{\partial^2 f}{\partial x^2}(t + c_i h, x_i) (\frac{\partial x_i}{\partial h})^2 + \frac{\partial f}{\partial x}(t + c_i h, x_i) \frac{\partial^2 x_i}{\partial h^2}) \\ \frac{\partial^2 x_i}{\partial h^2} &= \sum_{1 \leq j < i} a_{i,j} (c_j \frac{\partial f}{\partial t}(t + c_j h, x_j) + \frac{\partial f}{\partial x}(t + c_j h, x_j) \frac{\partial x_j}{\partial h}) + h \sum_{1 \leq j < i} a_{i,j} \dots \end{aligned}$$

Donc pour $h = 0$, on obtient :

$$\begin{aligned} \frac{\partial^2 x_i}{\partial h^2} |_{h=0} &= \sum_{1 \leq j < i} a_{i,j} (c_j \frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x) \frac{\partial x_j}{\partial h} |_{h=0}) \\ &= \sum_{1 \leq j < i} a_{i,j} c_j [\frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x) f(t, x)] \end{aligned}$$

et donc

$$\begin{aligned} \frac{\partial^2 F}{\partial h^2}(t, x, 0) &= \sum_{1 \leq i \leq q} b_i c_i^2 [\frac{\partial^2 f}{\partial t^2}(t, x) + 2 \frac{\partial^2 f}{\partial t \partial x}(t, x) f(t, x) \\ &\quad + \frac{\partial^2 f}{\partial x^2}(t, x) f(t, x)^2] + \sum_{1 \leq i \leq q} b_i \frac{\partial f}{\partial x}(t, x) \sum_{1 \leq j < i} a_{i,j} c_j [\frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x) f(t, x)] \end{aligned}$$

De l'autre côté, on calcule :

$$\begin{aligned} f^{[2]}(t, x) &= \frac{\partial f^{[1]}}{\partial t}(t, x) + \frac{\partial f^{[1]}}{\partial x}(t, x) f(t, x) \\ &= \dots = [\frac{\partial^2 f}{\partial t^2}(t, x) + 2 \frac{\partial^2 f}{\partial t \partial x}(t, x) f(t, x) \\ &\quad + \frac{\partial^2 f}{\partial x^2}(t, x) f(t, x)^2] + \frac{\partial f}{\partial x}(t, x) [\frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x) f(t, x)] \end{aligned}$$

En prenant les fonction $f(t, x) = t^2$ puis $f(t, x) = t + x$, on obtient aisément que la condition de la Proposition 7.1.16 :

$$\frac{\partial^2 F}{\partial h^2}(t, x, 0) = \frac{1}{3} f^{[2]}(t, x)$$

se traduit par les deux conditions :

$$\sum_{1 \leq i \leq q} b_i c_i^2 = \frac{1}{3}, \quad \sum_{1 \leq i \leq q, 1 \leq j < i} b_i a_{i,j} c_j = \frac{1}{6}$$

On poursuivrait de même le calcul de $\frac{\partial^3 F}{\partial h^3}(t, x, 0)$ pour obtenir l'ordre ≥ 3 . ■

Remarque. La méthode présentée dans l'exemple 7.3.3 est d'ordre 3.

Table des Matières

Chapitre 1	<i>Rappel des résultats fondamentaux de calcul différentiel</i>	
1.1	Différentiabilité	3
1.2	Théorème des Accroissements Finis	5
1.3	Théorèmes des fonctions implicites	6
1.4	Théorème d'inversion locale	7
1.5	Notations différentielles	8
Chapitre 2	<i>Méthodes d'approximation numérique des fonctions</i>	
2.1	Approximation d'une dérivée	9
2.2	Interpolation d'une fonction par un polynôme	16
2.3	Interpolation d'une fonction par un polynôme de degré 1	17
2.4	Interpolation d'une fonction par un polynôme de degré 2.	21
2.5	Interpolation d'une fonction par un polynôme de degré n	24
2.6	Stabilité numérique du procédé d'interpolation de Lagrange	33
Chapitre 3	<i>Méthodes d'approximation numérique des intégrales</i>	
3.1	Formules d'intégration à 1 point	41
3.2	Une formule d'intégration à 2 points	47
3.3	Méthodes générales de quadrature	49
Chapitre 4	<i>Méthodes d'approximation numérique des solutions d'équations</i>	
4.1	Principe des méthodes itératives	61
4.2	Utilisation du théorème du point fixe pour des fonctions réelles de classe C^1 ..	63
4.3	Méthode de Newton pour les fonctions réelles	65
4.4	Méthode de la sécante pour les fonctions réelles	67
4.5	Cas des fonctions vectorielles	69
Chapitre 5	<i>Equations différentielles, existence et unicité des solutions</i>	
5.1	Introduction	75
5.2	Principaux résultats dans le cas Lipschitzien	82
5.3	Principaux résultats dans le cas localement lipschitzien	88
5.4	Appendice sur le schéma d'Euler	95
5.5	Autres énoncés	99
Chapitre 6	<i>Equations différentielles linéaires</i>	
6.1	Equations différentielles linéaires sans deuxième membre	101
6.2	Equations différentielles linéaires avec deuxième membre	106
6.3	Equations différentielles linéaires à coefficients constants	108
6.4	Equations différentielles linéaires d'ordre m à coefficients constants	109
6.5	Wronskien d'un système de solutions	114
Chapitre 7	<i>Méthodes d'approximation numérique à un pas des équations différentielles.</i>	
7.1	Convergence des schémas numériques explicites	118
7.2	Equations de type gradient et schémas implicites	129
7.3	Méthodes de Runge-Kutta	131

Index des définitions

<p>Application différentiable 3</p> <p>Applications uniformément lipschitziennes 76</p> <p>Approximation de la dérivée seconde . . 15</p> <p>Caractérisation de la résolvante 94</p> <p>Condition initiale 70</p> <p>Consistance et stabilité des schémas explicites 106</p> <p>Continuité des différentielles des fonctions implicites 1 6</p> <p>Continuité des différentielles des fonctions implicites 2 6</p> <p>Convergence des polynômes d'interpolation d'une fonction analytique 31</p> <p>Convergence des schémas explicites . . 109</p> <p>Convergence des schémas implicites . . 117</p> <p>Convergence du schéma du point milieu 111</p> <p>Degrés de liberté 70</p> <p>Difféomorphisme 7</p> <p>Différence divisée d'ordre k 25</p> <p>Différence divisée d'ordre 1 17</p> <p>Différence divisée d'ordre 2 22</p> <p>Différences finies 10</p> <p>Différentiabilité des fonctions implicites 6</p> <p>Définition des schémas explicites 106</p> <p>Définition des schémas implicites 116</p> <p>Définition du schéma d'Euler 74</p> <p>Définition du schéma du point milieu 106</p> <p>Dérivation suivant un vecteur 3</p> <p>Equation différentielle d'ordre m 73</p> <p>Equation différentielle d'ordre 2 72</p> <p>Equations différentielles de type gradient 86</p> <p>Equations différentielles linéaires homogène 93</p> <p>Erreur au n-ième pas de temps 107</p> <p>Erreur d'interpolation 18</p> <p>Erreur d'interpolation d'ordre n 26</p> <p>Erreur d'interpolation d'ordre 2 23</p> <p>Erreur d'interpolation pour les polynômes de degré 1 19</p> <p>Erreur dans l'approximation d'une dérivée 12</p>	<p>Erreur dans l'approximation de la dérivée seconde 15</p> <p>Erreur dans la formule des trapèzes . . . 45</p> <p>Erreur dans la formule du point milieu 42</p> <p>Erreur dans la formules des rectangles . 39</p> <p>Erreur dans la méthode de quadrature élémentaire 50</p> <p>Existence de solutions locales. 82</p> <p>Existence de solutions maximales 84</p> <p>Existence des fonctions implicites 6</p> <p>Exponentielle d'applications linéaires . 98</p> <p>Fonctions de Liapounov 83</p> <p>Fonctions localement lipschitziennes . . 80</p> <p>Formule de Taylor 13</p> <p>Formule de Taylor avec reste intégral . 42</p> <p>Formule de Trotter 114</p> <p>Formule des rectangles 38</p> <p>Formule des trapèzes 44</p> <p>Formule du point milieu pour les intégrales 41</p> <p>Lemme de Gronwall, version différentielle 76</p> <p>Lemme de Gronwall, version discrète . . 77</p> <p>Majoration de l'erreur dans l'approximation centrée d'une dérivée 14</p> <p>Majoration de l'erreur dans l'approximation d'une dérivée 13</p> <p>Méthode de la sécante 63</p> <p>Méthode de Newton 61</p> <p>Méthode de quadrature 47</p> <p>Opérateur d'interpolation de Lagrange 33</p> <p>Polynôme d'interpolation de Newton . . 17</p> <p>Polynômes d'interpolation de Lagrange 18</p> <p>Polynômes d'interpolation de Lagrange de degré n 25</p> <p>Polynômes d'interpolation de Lagrange de degré 2 22</p> <p>Polynômes et points d'interpolation . . 16</p> <p>Problème de valeurs initiales 70</p> <p>Propriétés de la résolvante 94</p> <p>Résolvante 94</p> <p>Schéma convergent 107</p> <p>Schéma d'Euler implicite 116</p>
---	--

Second théorème de la moyenne	39	Théorème des accroissements finis	12
Solution d'une équation différentielle . .	69	Théorème des Accroissements Finis scalaire	5
Solution d'une équation différentielle linéai- re homogène	93	Théorème des Accroissements Finis vecto- riel	5
Solution maximale	82	Théorème des Accroissements Finis à val- eurs scalaires	5
Solutions locales	81	Théorème du point fixe	57
Stabilité des solutions locales	82	Wronskien	103
Théorème de Cauchy-Lipschitz	76		
Théorème de Rolle	12		