Université Paris - Sud

CENTRE D'ORSAY

C3 Analyse Numérique

- :: -

Interpolation, Quadratures, Equations Différentielles.

Cours professé par : M. R. TEMAM
rédigé par : MM. CORNIL et SCHEURER

(année 1970/71)

(Secrétariat de Mathématique 2ème cycle)

Université Paris - Sud

CENTRE D'ORSAY

C3 Analyse Numérique

- :: -

Interpolation, Quadratures, Equations Différentielles.

Cours professé par : M. R. TEMAM

rédigé par : MM. CORNIL et SCHEURER

(année 1970/71)

(Secrétariat de Mathématique 2ème cycle)

Table des matières.

CHAPITRE I - INTERPOLATION, DERIVATION ET INTEGRATION NUMERIQUES.

A - INTERPOLATION I.1

1. Interpolation polynomiale I.1

2. Formules d'erreur I.2

3. Formules de Newton I.5

B - DERIVATION NUMERIQUE I.10

C - INTEGRATION NUMERIQUE I.12

CHAPITRE II - INTEGRATION NUMERIQUE DES EQUATIONS DIFFERENTIELLES

A - THEOREMES D'EXISTENCES II.1

Problème de Cauchy des équations différentielles .. II.1

Rappel. Théorème d'Ascoli.......................... II.4

Rappel. Théorème du point fixe..................... II.8

B - PROLONGEMENT DES SOLUTIONS II.11

a. Solutions maximales II.11

b. Critère de prolongement II.13

C - RESOLUTION NUMERIQUE (Méthodes de différence finie)....... II.17

1. Les méthodes à un et plusieurs pas II.17

2. Méthodes à un pas II.20

3. Méthodes à plusieurs pas. Utilisation des formules de quadrature II.27

4. Etude générale des méthodes à plusieurs pas. Stabilité II.32

Méthode d'Adams-Bashforth II.44

Méthode d'Adams-Moulton II.45

Méthode de Nyström II.45

Méthode de Milne Simpson II.45

CHAPITRE I

INTERPOLATION. DERIVATION ET INTEGRATION NUMERIQUES.

A INTERPOLATION

§ 1. Interpolation polynomiale.

Le problème est le suivant : étant donnés une fonction numérique f définie continue sur un intervalle $[a,b]$ et des points $x_o < x_1 \dots < x_n$ de $[a,b]$, trouver un polynôme qui prend les mêmes valeurs que f aux points $x_o \dots x_n$.

Proposition 1. Soient $x_o < x_1 \dots < x_n$ des points de $[a,b]$ et $f_o \dots f_n$, $n+1$ nombres réels, il existe un polynôme unique de degré n, tel que :

$$P_n(x_i) = f_i \qquad 0 \leqslant i \leqslant n .$$

Démonstration.

(1) Existence et unicité de P_n.

Soit $P_n(x) = \sum_{j=o}^{n} a_j x^j$ le polynôme. Les inconnues sont les coefficients $a_o \dots a_n$. Par hypothèse P_n est tel que :

$$P_n(x_o) = f_o \;\dots\; P_n(x_n) = f_n .$$

Ce système d'équations s'écrit sous forme matricielle :

$$\begin{pmatrix} 1 & x_o & x_o^2 & \dots & x_o^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \begin{pmatrix} a_o \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f_o \\ f_1 \\ \vdots \\ f_n \end{pmatrix} \quad \text{soit} \quad A\,a = f \qquad (1.1)$$

Le déterminant de A est un déterminant de Wandermonde : le système (1.1) est donc un système de Cramer car :

$$\det \mathbf{A} = \prod_{i<j} (x_i - x_j) \neq 0$$

car tous les points x_i sont distincts. Le système (1.1) admet donc une solution unique.

(2) Forme explicite de P_n .

Introduisons les polynômes élémentaires suivants :

* $$\omega_n(x) = (x-x_o) \dots (x-x_n) = \prod_{i=o}^{n} (x-x_i)$$

donc $$\omega'_n(x_k) = \prod_{\substack{i=o \\ i\neq k}}^{n} (x_k-x_j)$$

* $$\ell_k^{(n)}(x) = \frac{\prod\limits_{\substack{i=o \\ i\neq k}}^{n} (x-x_i)}{\prod\limits_{i\neq k}^{n} (x_k-x_i)} = \frac{\omega_n(x)}{(x-x_k)\omega'_n(x_k)}$$

On vérifie que $$\ell_k^{(n)}(x_i) = \begin{cases} 0 & \text{si} \quad i \neq k \\ 1 & \text{si} \quad i = k \end{cases}$$

$$d^o \ \ell_k^{(n)}(x) = n \ .$$

Considérons alors le polynôme $$\boxed{P_n(x) = \sum_{k=o}^{n} f_k \frac{\omega_n(x)}{(x-x_k)\omega'_n(x_k)} = \sum_{k=o}^{n} f_k \ \ell_k^{(n)}(x).} \quad (1.2)$$

Ce polynôme est de degré n et d'autre part :

$$P_n(x_i) = \sum_{k=o}^{n} f_k \ \ell_k^{(n)}(x_i) = f_i \qquad 0 \leqslant i \leqslant n \ .$$

Il vérifie donc le système (1.1) ; comme ce dernier admet une solution unique, le polynôme (1.2) est bien la solution cherchée.

Définition. Polynôme d'interpolation.

Soit $f : [a,b] \to \mathbb{R}$ où f est continu. On note $P_n(f \ ; x)$ et on appelle polynôme d'interpolation de f aux points $x_o \dots x_n$ de $[a,b]$ l'unique polynôme P_n de degré n tel que :

$$P_n(f \ ; x_i) = f(x_i) = f_i \qquad 0 \leqslant i \leqslant n \ .$$

Remarque. Le polynôme $P_n(f \ ; x)$ "approche" la fonction, mais il serait faux d'en déduire que lorsque $n \to \infty$, P_n converge uniformément vers f dans $[a,b]$.

§ 2. Formules d'erreur.

Proposition 2.1. Soient $f \in \mathcal{C}^{n+1}([a,b])$ et $x_o \dots x_n$ des points de $[a,b]$.

Pour tout $x \in]a,b[$, il existe ξ_x tel que :

$$\operatorname{Min}(x,\ x_o \ \dots\ x_n) \leqslant \xi_x \leqslant \operatorname{Max}\ (x,\ x_o\ \dots\ x_n) \quad \text{et}$$

$$f(x) - P_n(f\ ;\ x) = \frac{(x-x_o)\dots(x-x_n)}{(n+1)!}\ f^{(n+1)}(\xi_x) = \frac{\omega_n(x)}{(n+1)!}\ f^{(n+1)}\ (\xi_x) \qquad (2.1)$$

<u>Démonstration</u>. Considérons la fonction de t :

$$\psi(t) = f(t) - P_n(f\ ;\ t) - \frac{(t-x_o)\dots(t-x_n)}{(x-x_o)\dots(x-x_n)}\ [f(x) - P_n(f;x)]\ .$$

Alors $\psi \in \mathscr{C}^{n+1}([a,b])$ et s'annule aux points $t = x_i$, $0 \leqslant i \leqslant n$ et $t = x$.

D'après le lemme ci-après il existe donc $\xi = \xi_x$ tel que :

$$\psi^{(n+1)}(\xi_x) = 0\ .$$

Donc il existe ξ_x tel que :

$$\psi^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - \frac{f(x)-P_n(f\ ;\ x)}{(x-x_o)\dots(x-x_n)}\ (n+1)! = 0\ .$$

(En effet $P_n(f\ ;\ t)$ étant de degré n $\dfrac{d^{n+1}}{dt^{n+1}}\ P_n(f\ ;\ t) = 0$ et $(t-x_o)\dots(t-x_n)$ étant également de degré n, $\dfrac{d^{n+1}}{dt^{n+1}}\ \omega_n(t) = (n+1)!\ \text{coef.}[t^{n+1}$ dans $\omega_n(t)] = (n+1)!)$.

<u>Lemme</u>. Soit $\psi : [a,b] \to \mathbb{R}$, $\psi \in \mathscr{C}^{n+1}[a,b]$,

(plus précisément f est n fois continument dérivable dans $[a,b]$ et sa dérivée $(n+1)^=$ existe dans $]a,b[$),

tel que $\psi(y_j) = 0$ avec $a < y_o < \dots < y_{n+1} < b$. Alors il existe ξ , $y_o < \xi < y_{n+1}$ tel que :

$$\psi^{(n+1)}(\xi) = 0\ .$$

<u>Démonstration</u>. Elle repose sur le théorème de Rolle

$$\begin{array}{llllll}
\downarrow\text{Rolle} & \psi = 0 & \text{en} & y_o \quad y_1 & \dots\dots & y_n \quad y_{n+1} \quad \text{soit } n+2 \text{ points} \\
 & \psi' = 0 & \text{en} & y_o^{(1)} & \dots\dots & y_n^{(1)} \quad \text{soit } n+1 \text{ points} \\
\text{Rolle}\downarrow & \psi'' = 0 & \text{en} & y_o^{(2)} & \dots\dots & y_{n-1}^{(2)} \quad \text{soit } n \text{ points} \\
 & \vdots & & & & \\
 & \psi^{(n+1)} = 0 & \text{en} & & y_o^{(n+1)} = \xi & \quad \text{soit } 1 \text{ point.}
\end{array}$$

<u>Proposition</u> 2.2. Si f est analytique dans un ouvert Ω du plan complexe, qui contient l'intervalle réel $[a,b]$. Si Γ est une courbe de Jordan régulière contenant $[a,b]$ et contenu dans Ω alors :

$$|f(x) - P_n(f\ ;\ x)| \leqslant \frac{\ell\mu}{2\pi\delta^{n+2}}\ |x-x_o|\ldots|x-x_n| \tag{2.2}$$

où ℓ = longueur de Γ, $\mu = \max\limits_{x\in\Gamma} |f(x)|$ $\delta = d([a,b], \Gamma)$.

<u>Démonstration</u>. On utilise la formule de Cauchy :

$$j(x\ ;\ \Gamma)\ f^{(n+1)}(\xi) = \frac{(n+1)!}{2\pi i} \int_\Gamma \frac{f(z)}{(z-\xi)^{n+2}}\, dz\ .$$

D'après la formule (2.1)

$$|f(x) - P_n(f\ ;\ x)| = \frac{|x-x_o|\ldots|x-x_n|}{(n+1)!}\, |f^{(n+1)}(\xi)|\ .$$

Comme $j(x\ ;\ \Gamma) = 1$:

$$|f^{(n+1)}(\xi)| = \frac{(n+1)!}{2\pi}\left|\int_\Gamma \frac{f(z)}{(z-\xi)^{n+2}}\, dz\right| \leqslant \frac{(n+1)!}{2\pi}\ \frac{\mu\ell}{\delta^{n+2}}$$

donc

$$\begin{aligned}|f(x) - P_n(f\ ;\ x)| &\leqslant \frac{|x-x_o|\ldots|x-x_n|}{(n+1)!}\ \frac{(n+1)!}{2\pi}\ \frac{\mu\ell}{\delta^{n+2}}\\ &\leqslant \frac{|x-x_o|\ldots|x-x_n|}{2\pi}\ \frac{\mu\ell}{\delta^{n+2}}\end{aligned}$$

On a défini P_n comme un opérateur dans $\mathscr{C}^{(n+1)}[a,b]$. Définissons l'opérateur L comme suit

$$L : f \in \mathscr{C}^{n+1}[a,b] \mapsto f - P_n(f) = L(f)\ .$$

On a alors la :

<u>Proposition</u> 2.3. <u>Formule d'erreur de Peano</u>.

Pour $f \in \mathscr{C}^{n+1}[a,b]$ on a : $L(f)(x) = f(x) - P_n(f\ ;\ x) = \int_a^b f^{(n+1)}(t)k(x,t)dt$ (2.3)

où

$$k(x,t) = \frac{1}{n!}\ L_x\ [(x-t)_+^n]$$

L_x désignant l'opérateur L appliqué a : $x \mapsto (x-t)_+^n = \begin{cases}(x-t)^n & \text{si } (x-t) \geqslant 0\\ 0 & \text{si } (x-t) < 0\end{cases}$.

Démonstration. Utilisons la formule de Taylor à reste intégrale :

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

donc $L(f)(x) = f(x) - P_n(f\ ;\ x) = \frac{1}{n!}\int_a^b f^{(n+1)}(t) L_x (x-t)_+^n\, dt$

$$= \int_a^b f^{(n+1)}(t)\ k(x,t)\ dt$$

Proposition 2.4. Pour $f \in \mathscr{C}^{n+1}[a,b]$ on a :

$$L(f)(x) = \sum_{i=o}^{i=n} \frac{\ell_i^{(n)}(x)}{n!} \int_{x_i}^{x} (x_i-t)^n\, f^{(n+1)}(t)dt \quad , \quad x_o \dots x_n \in [a,b] .$$

Démonstration. Par définition de L :

$$k(x,t) = \frac{1}{n!}\ L_x[(x-t)_+^n] = \frac{1}{n!}\ (x-t)_+^n - \frac{1}{n!}\sum_{i=o}^{n} (x_i-t)_+^n \cdot \ell_i^{(n)}(x)$$

donc d'après la formule (2.3) :

$$L(f)(x) = \int_a^b f^{(n+1)}(t)k(x,t)dt = \int_a^b f^{(n+1)}(t)[\frac{1}{n!}(x-t)_+^n - \sum_{i=o}^{n}(x_i-t)_+^n\ \ell_i^{(n)}(x)]dt$$

$$= \frac{1}{n!}\int_a^b f^{(n+1)}(t)(x-t)_+^n dt - \frac{1}{n!}\sum_{i=o}^{n}\int_a^b f^{(n+1)}(t)(x_i-t)_+^n\ \ell_i^{(n)}(x)\ dt$$

$$= \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n dt - \sum_{i=o}^{n} \frac{\ell_i^{(n)}(x)}{n!}\int_a^{x_i} f^{(n+1)}(t)\ (x_i-t)^n\ dt .$$

D'après la proposition 1 $(x-t)^n$, étant de degré n , est son propre polynôme d'interpolation : $(x-t)^n = P_n((x-t)^n, x) = \sum_{i=o}^{n} (x_i-t)^n\ \ell_i^{(n)}(x)$, donc

$$L(f)(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t) \sum_{i=o}^{n}(x_i-t)^n \ell_i^{(n)}(x)dt + \sum_{i=o}^{n} \frac{\ell_i^{(n)}(x)}{n!}\int_{x_i}^{a} f^{(n+1)}(t)(x_i-t)^n dt$$

$$= \sum_{i=o}^{n} \frac{\ell_i^{(n)}(x)}{n!}\int_{x_i}^{x} f^{(n+1)}(t)(x_i-t)^n\ dt$$

§ 3. Formules de Newton.

Soient $f \in \mathscr{C}^{n+1}[a,b]$ et $a < x_o \dots < x_n < b$. Cherchons une représentation du polynôme d'interpolation P_n du type $\sum_{k=o}^{n} a_k\, \omega_{k-1} = P_n$.

Une telle représentation existe puisque les $(n+2)$ polynômes élémentaires $\omega_{-1}(x) = 1$, $\omega_o(x) = x-x_o \dots \omega_k(x) = \prod_{i=o}^{k} (x-x_i) \dots \omega_n(x) = \prod_{i=o}^{n} (x-x_i)$ forment une base de l'espace vectoriel $\mathcal{P}_{n+1}$ des polynômes de degré $\leqslant n+1$ [rappelons que $\dim \mathcal{P}_{n+1} = n+2$]. En effet : $\omega_k(x) = x^k + v_{k-1}^{(k)} x^{k-1} + \dots + v_o^{(k)}$. Donc si on écrit $\omega_{-1}, \omega_o, \dots, \omega_k, \dots, \omega_n$ dans la base $1, x, \dots, x^{n+1}$ de $\mathcal{P}_{n+1}$ on obtient la matrice de changement de base :

$$\mathcal{M} = \begin{array}{c} \begin{array}{cccccc} \omega_{-1} & \omega_o & & \omega_k & & \omega_n \end{array} \\ \begin{pmatrix} 1 & -x_o & & v_o^{(k)} & & v_o^{(n)} \\ 0 & 1 & & \vdots & & \vdots \\ \vdots & & \ddots & \vdots & & v_i^{(n)} \\ \vdots & & & 1 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & & & & & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ x \\ x^{i+1} \\ x^{k+1} \\ \\ x^{n+1} \end{array}$$

et det $\mathcal{M} = +1$.

Soit P_j le polynôme $\sum_{k=o}^{j} a_k\, \omega_{k-1}$. On a $d^o\, P_j = j$.

On vérifie aisément par récurrence descendante que :

$$P_n(x_i) = f_i \qquad 0 \leqslant i \leqslant n$$
$$\dots\dots\dots\dots$$
$$P_j(x_i) = f_i \qquad 0 \leqslant j \leqslant i$$
$$\dots\dots\dots\dots$$

De sorte que $P_n \dots P_j \dots$ sont respectivement les polynômes d'interpolation de f aux points $x_o \dots x_n, \dots, x_o \dots x_j, \dots$. Or on peut écrire :

$$P_n(f\;;x) = \sum_{k=o}^{n} \ell_k^{(n)}(x) f_k = \sum_{k=o}^{n} \frac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)}\, f_k$$
$$= \sum_{k=o}^{n} f_k\, \frac{(x-x_o)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{\omega_n'(x_k)}.$$

D'où $P_n(f\;;x) = x^n \sum_{k=o}^{n} \frac{f_k}{\omega_n'(x_k)} + x^{n-1}(\quad) + \dots$.

Identifions cette dernière égalité avec :

$$P_n(f\;;x) = \sum_{k=o}^{n} a_k\, \omega_{k-1}(x) = a_n\, \omega_{n-1}(x) + a_{n-1}\, \omega_{n-1}(x) + \dots$$
$$= a_n[x^n + (\dots)x^{n-1} + \dots] + a_{n-1}[x^{n-1} + (\quad)\, x^{n-2} + \dots] + \dots$$
$$= a_n\, x^n + \dots$$

On a alors : $a_n = \sum_{k=o}^{n} \frac{f_k}{\omega'_n(x_k)} = \sum_{k=o}^{n} \frac{f(x_k)}{\omega'_n(x_k)}$.

De même a_j $(j < n)$ coefficient de x^j dans $P_j(f\ ;\ x) = \sum_{o}^{j} a_k\, \omega_{k-1}(x)$ est égal par identification avec $P_j(f\ ;\ x) = \sum_{k=o}^{j} \ell_k^{(j)}(x)\, f_k$ à $\sum_{k=o}^{j} \frac{f_k}{\omega'_j(x_k)}$.

Ainsi les coefficients a_k de $P_n = \sum_{k=o}^{n} a_k\, \omega_{k-1}$ sont donnés par :

$$\boxed{a_k = \sum_{i=o}^{k} \frac{f(x_i)}{\omega'_k(x_i)} = \sum_{i=o}^{k} \frac{f(x_i)}{\prod\limits_{\substack{\ell=o \\ \ell\neq i}}^{k} (x_i - x_\ell)} \qquad 0 \leqslant k \leqslant n} \tag{3.1}$$

Définition. Le coefficient a_k s'appelle la $k^{=}$ différence divisée de f aux points $x_o \dots x_k$ et on note : $a_k = f[x_o \dots x_k]$.

Résumons les résultats précédents. On a la :

Proposition 3.1. Soient $f \in \mathscr{C}[a,b]$ et $x_o \dots x_n$ des points de $[a,b]$ alors le polynôme d'interpolation de f aux points $x_o \dots x_n$ s'écrit :

$$P_n(f\ ;\ x) = \sum_{k=o}^{n} f[x_o,\dots, x_k]\, \omega_{k-1}(x) \ .$$

Proposition 3.2. On a la relation de récurrence suivante entre les différences divisées : $f[x_o] = f(x_o) = f_o$, et pour $n \geqslant 1$

$$f[x_o,\dots, x_n] = \frac{f[x_1,\dots, x_k] - f[x_o,\dots, x_{n-1}]}{x_n - x_o} \ .$$

Démonstration.

$$\sum_{k=o}^{n} \frac{f_k}{\omega'_n(x_k)} (x_n - x_o) = a_n(x-x_o) = \sum_{k=o}^{n} \frac{f_k}{\omega'_n(x_k)} (x_n - x_k) + \sum_{k=o}^{n} \frac{f_k}{\omega'_r(x_k)} (x_k - x_o)$$

$$= \sum_{k=o}^{n-1} \frac{f_k}{\omega'_n(x_k)} + \sum_{k=1}^{n} \frac{f_k}{\omega'_n(x_k)}$$

$$= - f[x_o,\dots, x_{n-1}] + f[x_1,\dots, x_n]$$

d'où $$a_n = \frac{f[x_1,\dots, x_n] - f[x_o,\dots, x_{n-1}]}{x_n - x_o}$$

Nous allons préciser l'expression de $P_n(f\ ;\ x) = \sum_{k=o}^{n} a_k\, \omega_{k-1}(x)$ dans deux cas particuliers.

* Cas des points équidistants.

Définissons les points $x_o, \dots, x_n$ de l'intervalle $[a,b]$ de la façon suivante : $x_j = x_o + j\,h$, $0 \leqslant j \leqslant n$, h étant le pas de l'intervalle.

Proposition 3.3. Soient $f \in \mathcal{C}[a,b]$, $x_j = x_o + j\,h$, $0 < j < n$ des points de $[a,b]$, alors $P_n(f\,;\,x) = \sum_{k=o}^{n} a_k\,\omega_{k-1}(x)$ avec

$$a_k = \sum_{i=o}^{k} \frac{(-1)^{k-i}}{h^k\,k!} \binom{k}{i} f_i \,.$$

Démonstration. D'après la formule (3.2)

$$a_k = \sum_{i=o}^{k} \frac{f_i}{\omega_k(x_i)} = \sum_{i=o}^{k} \frac{f_i}{(x_i-x_o)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_k)} \,.$$

Mais $x_j = x_o + j\,h$, $0 \leqslant j \leqslant n$. Donc en remplaçant dans l'expression ci-dessus

$$a_k = \sum_{i=o}^{k} \frac{(-1)^{k-i}}{h^k\,i!\,(k-i)!} f_i = \sum_{i=o}^{k} \frac{(-1)^{k-i}}{h^k\,k!} \binom{k}{i} f_i$$

* Cas des différences progressives.

Considérons une suite infinie $(y) = y_o, y_1, \dots, y_k, \dots$. On définit dans l'espace vectoriel des suites, l'opérateur linéaire ∇ de différence progressive par : $(\nabla y)_k = \nabla\{y_k\}_{k\in\mathbb{N}} = \{y_{k+1} - y_k\}_{k\in\mathbb{N}}$. On peut aussi définir les puissances ∇^n de cet opérateur. En effet on a le :

Lemme. $(\nabla^n y)_k = \sum_{j=o}^{n} \left[(-1)^{n-j} \binom{n}{j} y_{k+j}\right]$.

Démonstration. Raisonnons par récurrence.

* Par définition de ∇y_k la proposition est vraie pour $n = 1$.

* Admettons le résultat pour n :

$$(\nabla^n y)_k = \sum_{j=o}^{n} \left[(-1)^{n-j} \binom{n}{j} y_{k+j}\right] .$$

* Démontrons le pour $n+1$:

$$
\begin{aligned}
(\nabla^{n+1}y)_k &= (\nabla^n y)_{k+1} - (\nabla^n y)_k \\
&= \sum_{j=0}^{n} (-1)^{n-j}\binom{n}{j} y_{k+1+j} - \sum_{j=0}^{n} (-1)^{n-j}\binom{n}{j} y_{k+j} \\
&= \sum_{j=1}^{n+1} (-1)^{n+1-j}\binom{n}{j+1} y_{k+j} - \sum_{j=0}^{n} (-1)^{n-j}\binom{n}{j} y_{k+j} \\
&= \sum_{j=1}^{n} (-1)^{n+1-j} y_{k+j}\left(\binom{n}{j-1}+\binom{n}{j}\right)+(-1)^{o}\binom{n}{n}+(-1)^{k+1}\binom{n}{o} \\
&= \sum_{j=1}^{n} (-1)^{n+1-j} y_{k+j}\binom{n+1}{j}+(-1)^{o}\binom{n+1}{n+1}+(-1)^{n+1}\binom{n+1}{o} \\
&= \sum_{j=o}^{n+1} (-1)^{n+1-j} y_{k+j}\binom{n+1}{j}
\end{aligned}
$$

Exprimons alors le polynôme $P_n(f\ ;\ x)$ à l'aide des différences progressives.

<u>Proposition</u> 3.4. Soit $f \in \mathcal{C}[a,b]$, $x_o,\ldots, x_n$ des points de $[a,b]$ tels que $x_j = x_o + j\,h$, $0 \leqslant j \leqslant n$ et posons $s = \frac{x-x_o}{h}$. Alors

$$P_n(f\ ;\ x) = \sum_{k=o}^{n} \binom{s}{k} \nabla^k f_o \ .$$

<u>Démonstration</u>. D'après la proposition 3.3. on a vu que :

$$a_k = \sum_{i=o}^{k} \frac{(-1)^{k-i}}{h^k\ k!} \binom{k}{i} f_i$$

ce qui s'écrit encore d'après le lemme ci-dessus :

$$a_k = \frac{1}{h^k\ k!} (\nabla^k f)_j = \frac{1}{h^k\ k!} \nabla^k f_o \ .$$

On en déduit donc que :

$$
\begin{aligned}
P_n(f\ ;\ x) &= \sum_{k=o}^{n} a_k\, \omega_{k-1}(x) = \sum_{k=o}^{n} a_k (x-x_o)\ldots(x-x_j)\ldots(x-x_{k-1}) \\
&= \sum_{k=o}^{n} \frac{1}{h^k\ k!} \nabla^k f_o\ (x-x_o)\ldots(x-x_j)\ldots(x-x_{k-1}) \\
&= \sum_{k=o}^{n} \frac{1}{h^k\ k!} \nabla^k f_o\ (x-x_o)\ldots(x-x_o-j\,h)\ldots(x-x_o-(k-1)h)
\end{aligned}
$$

car $x_j = x_o + j\,h$, $0 \leqslant j \leqslant n$. Effectuons le changement de variable $s = \frac{x-x_o}{h}$:

$$
\begin{aligned}
P_n(f\ ;\ x) &= \sum_{k=o}^{n} \frac{1}{h^k\ k!} \nabla^k f_o\ sh(s-1)h\ \ldots\ (s-(k-1))h \\
&= \sum_{k=o}^{n} \frac{s(s-1)\ldots(s-(k-1))}{k!} \nabla^k f_o = \sum_{k=o}^{n} \binom{s}{k} \nabla^k f_o
\end{aligned}
$$

B DERIVATION NUMERIQUE

1^e approximation.

$$f'(x) \approx \frac{f(x+\Delta x)-f(x)}{\Delta x} .$$

C'est l'approximation la plus simple, mais elle est toujours très mauvaise (perte de chiffres significatifs).

2^e approximation.

Cette approximation, valable lorsqu'on connait f en plusieurs points $x_o, \ldots, x_n$ "voisins" de x, est plus raffinée. Elle consiste à approcher la dérivée de f au point x par la dérivée du polynôme d'interpolation de f aux points $x_o, \ldots, x_n$:

$$f'(x) \approx P_n'(f \; ; \; x).$$

Proposition 4. Soit f $(n+1)$ fois continument différentiable sur $[a,b]$. Soit $Sp(x_o, \ldots, x_n, x)$ l'intervalle engendré par $x_o, \ldots, x_n, x$. Alors il existe ξ_x tel que :

$$\text{(i)} \quad f'(x_j) - P_n'(f \; ; \; x_j) = \frac{1}{(n+1)!} \, \omega_n'(x_j) \, f^{(n+1)}(\xi_x)$$

et on a
$$\text{(ii)} \quad f'(x_j) - P_n'(f \; ; \; x_j) = \sum_{i=0}^{n} \left(\frac{\ell_i^{(x)}(x_j)}{n!} \right)' \int_{x_i}^{x_j} (x_i - t)^n \, f^{(n+1)}(t) dt.$$

Démonstration. (i) Sous l'hypothèse ci-dessus, on a vu que (prop. 2.1) il existe ξ_x tel que

$$(2.1) \quad \mathrm{Min}(x_o, \ldots, x_n, x) \leqslant \xi_x \leqslant \mathrm{Max}(x_o, \ldots, x_n, x) \quad \text{et} \quad f(x) - P_n(f;x) = \frac{\omega_n(x)}{n!} f^{(n+1)}(\xi_x).$$

Le problème est de savoir si la fonction $x \to f^{(n+1)}(\xi_x)$ est continument dérivable. Posons :

$$g(x) = f^{(n+1)}(\xi_x) = \frac{f(x) - P_n(f\ ;\ x)}{\omega_n(x)} (n+1)! \quad \text{si} \quad x \neq x_j$$

$$g(x_j) = \frac{f'(x_j) - P_n'(f\ ;\ x_j)}{\omega_n'(x_j)} (n+1)\ !\ , \quad 0 \leqslant j \leqslant n\ .$$

On vérifie que g est continument dériable sur $]a,b[$. $\frac{d}{dx}[g(x)]_{x=x_j}$ a donc un sens. On peut donc dériver la formule d'erreur (2.1) et faire $x = x_j$. On obtient :

$$\left[\frac{d}{dx}(f(x) - P_n(f\ ;\ x))\right]_{x=x_j} = \frac{\omega_n'(x_j)}{(n+1)!} f^{(n+1)}(\xi_x) + \frac{\omega_n(x_j)g'(x_j)}{(n+1)!}$$

or par définition $\omega_n(x_j) = 0$. Donc :

$$f'(x_j) - P_n'(f\ ;\ x_j) = \frac{\omega_n'(x_j)}{(n+1)!} f^{(n+1)}(\xi_x)$$

■

Il est plus rapide d'appliquer le théorème de Rolle au polynôme

$$[f'(x_j) - P_n'(f\ ;\ x_j)] \prod_{i=1}^{n} (u - x_i) - \prod_{i \neq j} (x_j - x_i)[f(u) - P_n(f,u)]$$

qui possède x_j comme racine double et les x_i comme racine simple.

(ii) Sous l'hypothèse ci-dessus, on a vu que (prop. 2.4) :

$$f(x) - P_n(f\ ;\ x) = \sum_{i=0}^{n} \frac{\ell_i^{(n)}(x)}{n!} \int_{x_i}^{x} (x_i - t)^n f^{(n+1)}(t)dt\ .$$

Dérivons cette égalité et faison $x = x_j$; on obtient (la dérivation étant licite) :

$$f'(x_j) - P_n'(f\ ;\ x_j) = \sum_{i=0}^{n} \frac{(\ell_i^{(n)}(x_j))'}{n!} \int_{x_i}^{x} (x_i - t)^n f^{(n+1)}(t)dt$$

$$+ \underbrace{\sum_{i=0}^{n} \ell_i^{(n)}(x)\ (x_o - x_j)^n f^{(n+1)}(x_j)}_{\text{terme nul}}\ .$$

D'où :

$$f'(x_j) - P_n'(f\ ;\ x_j) = \sum_{i=0}^{n} \frac{(\ell_i^{(n)}(x_j))'}{n!} \int_{x_i}^{x_j} (x_i - t)^n f^{(n+1)}(t)\ dt$$

3^e approximation.

On a vu (prop. 3.4) qu'on pouvait écrire :

$$P_n(f\ ;\ x) = \sum_{k=0}^{n} \binom{s}{k} \nabla^k f_0 \quad \text{avec} \quad s = \frac{x-x_0}{h} .$$

Par suite :

$$P'_n(f\ ;\ x) = \frac{1}{h} \frac{d}{ds} \sum_{k=0}^{n} \binom{s}{k} \nabla^k f_0 .$$

Donc : $$P'_n(f\ ;\ x_0) = \sum_{k=0}^{n} \frac{1}{h} \frac{(-1)^{k-1}}{k} \nabla^k f_0 .$$

Par conséquent, on peut écrire :

$$f'(x_0) \approx P'_n(f\ ;\ x_0) = \sum_{k=0}^{n} \frac{(-1)^{k-1}}{k\ h} \nabla^k f_0$$

C- INTEGRATION NUMERIQUE

Soit $x \mapsto \mu(x)$ une fonction "poids" que l'on prendra dans $L^1([a,b])$. On veut calculer de manière approchée $\int_a^b f(x)\mu(x)dx$, connaissant f en certains points $x_0,\ldots, x_n$ de $[a,b]$.

<u>Définition</u> 5.1. On appelle <u>formule de quadrature</u> tout formule d'approximation du genre :

$$\int_a^b f(x)\mu(x)dx \approx \sum_{j=0}^{n} A_j^{(n)} f_j$$

où $f_j = f(x_j)$, $0 \leqslant j \leqslant n$.

Ex : Formule des trapèzes.

$x \mapsto \mu(x) = 1$ $\qquad A_0^{(n)} = A_n^{(n)} = \frac{h}{2}$

$x_0 = a$ $\qquad x_j = x_0 + j\,h$, $0 \leqslant j \leqslant n$

$x_n = b$ $\qquad A_j^{(n)} = h$, $1 \leqslant j \leqslant n-1$

$$\int_a^b f(x)dx = h\left[\frac{f_0}{2} + f_1 + \ldots + f_{n-1} + \frac{f_n}{2}\right] .$$

(1) Une <u>1^e^ méthode</u> consiste à approcher la fonction par un polynôme, puis à calculer l'intégrale de ce polynôme multiplié par $\mu(x)$.

<u>Définition</u> 5.2. On appelle <u>formule de quadrature du type interpolation</u>, toute approximation du type :

$$\int_a^b f(x)\mu(x)dx \approx \int_a^b P_n(f\ ;\ x)\mu(x)dx$$

où $P_n(f\ ;\ x)$ est le polynôme d'interpolation de f aux points $x_o, \dots, x_n$ de $[a,b]$: $P_n(f\ ;\ x) = \sum_{j=o}^{n} \ell_j^{(n)}(x)\ f(x_j)$. On a donc explicitement :

$$\int_a^b f(x)\mu(x)dx \approx \sum_{j=o}^{n} f(x_j) \int_a^b \ell_j^{(x)}(x)\mu(x)dx\ .$$

On pose

$$\boxed{C_j^{(n)} = \int_a^b \ell_j^{(n)}(x)\mu(x)dx}\ .$$

Dans les applications les $C_j^{(n)}$ doivent être connus explicitement. On va donc préciser leur valeur. On a vu que :

$$\ell_j^{(n)} = \frac{\omega_n(x)}{(x-x_j)\omega_n'(x_j)}\ .$$

Donc :

$$C_j^{(n)} = \int_a^b \ell_j^{(n)}(x)\mu(x)dx = \int_a^b \frac{\omega_n(x)}{(x-x_j)\omega_n'(x_j)}\ \mu(x)\ dx$$

$$= \int_a^b \frac{(x-x_o)\dots(x-x_k)\dots(x-x_n)}{(x-x_j)(x_j-x_o)\dots(x_j-x_{j-1})(x_j-x_{j+1})\dots(x_j-x_n)}\ \mu(x)dx\ .$$

* <u>Supposons que les points</u> $x_o, \dots, x_n$ <u>soient équidistants</u>. Autrement dit :

$x_o = a$

$x_n = b \quad , \quad x_j = x_o + j\ h \quad , \quad 0 \leqslant j \leqslant n$. Posons $s = \dfrac{x-x_o}{h}$.

Alors

$$\ell_j^{(n)} = \frac{(x-x_o)\dots(x-x_{j-1})(x-x_{j+1})\dots(x-x_n)}{(x_j-x_o)\dots(x_j-x_{j-1})(x_j-x_{j+1})\dots(x_j-x_n)} = \frac{(x-x_o)\dots(x-x_{j-1})(x-x_{j+1})\dots(x-x_n)}{jh(j-1)h\dots h(-h)(-2h)\dots(-(n-j)h)}$$

$$= \frac{(x-x_o)\dots(x-x_{j-1})(x-x_{j+1})\dots(x-x_n)}{j!\ h^n(-1)^{n-j}(n-j)!} = \frac{s(s-1)\dots(s-j+1)(s-j-1)\dots(s-n)}{j!\ (-1)^{n-j}(n-j)!}$$

Donc $$C_j^{(n)} = \int_o^n \frac{s(s-1)\dots(s-j+1)(s-j-1)\dots(s-n)}{j!(-1)^{n-j}(n-j)!}\ \mu(sh + x_o)ds$$

$$= (-1)^{n-j}\ h \int_o^n \prod_{\substack{i=o\\ i\neq j}}^{n} (s-i)\ \frac{\mu(sh + x_o)}{j!(n-j)!}\ ds \quad .$$

Dans le cas où $x \mapsto \mu(x) = 1$, on obtient la formule de Newton-Cotes

$$\boxed{\overline{C}_j^{(n)} = \frac{1}{h}\ C_j^{(n)} = \frac{(-1)^{n-j}}{j!(n-j)!} \int_o^n \prod_{\substack{i=o\\ i\neq j}}^{n} (s-i)ds}$$

Définition 5.3. On dit qu'une formule de quadrature est exacte sur un espace de fonction $\mathcal{E}$ si :

$$\int_a^b f(x)\mu(x)dx = \sum_{j=c}^{n} A_j^{(n)}\ f(x_j) \qquad \forall\ f \in \mathcal{E}\ .$$

Proposition 5.4. Une formule de quadrature est exacte sur $\mathcal{P}_n$ (espace des polynômes de degré $\leqslant n$) si et seulement si elle est du type interpolation.

Démonstration.

* Supposons la formule du type interpolation. Alors :

$$\int_a^b f(x)\mu(x)dx \approx \sum_{j=o}^{n} \left(\int_a^b \ell_j^{(n)}(x)\mu(x)dx\right)\ f(x_j).$$

Mais par hypothèse $\mathcal{E} = \mathcal{P}_n$ donc $f = P_n(f)$ $\forall\ f \in \mathcal{P}_n$. Donc : $\forall f \in \mathcal{P}_n$

$$\int_a^b f(x)\mu(x)dx = \int_a^b P_n(f\ ;\ x)\mu(x)dx$$

$$= \int_a^b \sum_{j=o}^{n} \ell_j^{(n)}(x)\mu(x)dx)f(x_j)$$

$$= \sum_{j=o}^{n} \left(\int_a^b \ell_j^{(n)}(x)\mu(x)dx\right)\ f(x_j)$$

$$= \sum_{j=o}^{n} A_j^{(n)}\ f(x_j)\ .$$

La formule est donc exacte sur $\mathcal{P}_n$.

* Inversement, supposons la formule de quadrature exacte sur $\mathcal{P}_n$. Elle l'est donc en particulier pour les polynômes élémentaires $\ell_i^{(n)}$, $0 \leqslant i \leqslant n$, d'où :

$$\int_a^b \ell_i^{(n)}(x)\mu(x)dx = \sum_{j=o}^{n} A_j^{(n)}\ \ell_i^{(n)}(x_j) = A_i^{(n)} \qquad \forall i\ ,\ 0 \leqslant i \leqslant r\ ,$$

donc $$\sum_{j=o}^{n} A_j^{(n)} f(x_j) = \sum_{j=o}^{n} \left(\int_a^b \ell_j^{(n)}(x)\mu(x)dx\right) f(x_j)$$

$$= \int_a^b \left(\sum_{j=o}^{n} \ell_j^{(n)}(x)\ f(x_j)\right)\mu(x)dx = \int_a^b P_n(f\ ;\ x)\mu(x)dx$$

<u>Proposition</u> 5.5. <u>Formule</u> <u>d'erreur</u>.

Soit $f \in \mathcal{C}^{n+1}[a,b]$.

Si $$e_f = \int_a^b f(x)\mu(x)dx - \int_a^b P_n(f\ ;\ x)\ \mu(x)dx$$

alors $$|e_f| \leqslant \frac{M_{n+1}}{(n+1)!} \int_a^b |\omega_n(x)|\,|\mu(x)|\,dx$$

où $$M_{n+1} = t \in [a,b]\ .$$

<u>Démonstration</u>. On a vu que (prop. 2.1), il existait ξ_x tel que :

$$\mathrm{Min}(x_o,\ldots,\ x_n,\ x) \leqslant \xi_x \leqslant \mathrm{Max}(x_o,\ldots,\ x_n,\ x)$$

et : $$f(x) - P_n(f\ ;\ x) = \frac{\omega_n(x)}{(n+1)!}\ f^{(n+1)}(\xi_x)\ .$$

Par conséquent :

$$e_f = \int_a^b f(x)\mu(x)dx - \int_a^b P_n(f\ ;\ x)\mu(x)dx = \int_a^b [f(x)-P_n(f\ ;\ x)]\mu(x)dx$$

d'où $$|e_f| = \left|\int_a^b [f(x)-P_n(f\ ;\ x)]\mu(x)dx\right| \leqslant \int_a^b |f(x)-P_n(f\ ;\ x)|\,|\mu(x)|\,dx$$

soit $$|e_f| \leqslant \int_a^b \frac{|\omega_n(x)|}{(n+1)!}|f^{(n+1)}(\xi_x)|\,|\mu(x)|\,dx$$

$$\leqslant \sup_{t \in [a,b]} |f^{(n+1)}(t)| \int_a^b \frac{|\omega_n(x)|}{(n+1)!}\ |\mu(x)|\,dx$$

d'où $\quad |e_f| \leqslant \frac{M_{n+1}}{(n+1)!} \int_a^b |\omega_n(x)| \; |\mu(x)| dx$.

<u>Remarque</u>. * Soit f une fonction continue quelconque. L'erreur ne tend pas vers 0 , pour $n \to \infty$. D'autre part à partir de $n = 5, 6$ les $C_j^{(n)}$ sont de signes variables, ce qui introduit des instabilités numériques et limite considérablement les possibilités de la formule de Newton-Cotes.

* Utilisation pratique de Newton-Cotes.

(1) On subdivise l'intervalle $[a,b]$ en $x_o = a, x_1, \ldots, x_n = b$.

(2) Sur chaque intervalle on utilise une formule de quadrature du type interpolation avec n "petit" $(n \leqslant 4)$.

On retrouve comme cas particulier la formule de Simpson (cf. Ex.)

2 Une <u>2^e méthode</u> consiste à chercher s'il existe un choix judicieux des points $x_o, \ldots, x_n$ ($x_o, \ldots, x_n$, $(n+1)$ points de $[a,b]$) tel qu'une formule de quadrature basée sur $x_o, \ldots, x_n$ soit exacte sur un espace de polynômes $\mathcal{P}_{n'}$ avec $n' > n$. Les propositions 5.6 et 5.7 répondent complètement à la question dans le cas d'un poids $\mu > 0$. La réponse est alors liée aux polynômes orthogonaux relatifs à μ .

<u>Proposition</u> 5.6. Une formule de quadrature est exacte sur $\mathcal{P}_{2n+1}$ si et seulement si : (i) elle est exacte sur $\mathcal{P}_n \iff$ elle est du type interpolation

(ii) le polynôme $\omega_n(x) = (x-x_o)\ldots(x-x_n)$ vérifie :

$$\int_a^b x^p \omega_n(x)\mu(x)dx = 0 \quad , \quad 0 \leqslant p \leqslant n \,.$$

[autrement dit $\omega_n(x)$ est orthogonal pour le produit scalaire de poids μ à $1, x, \ldots, x^n$] .

Démonstration.

* Supposons la formule exacte sur $\mathcal{P}_{2n+1}$. Elle l'est donc à fortiori sur $\mathcal{P}_n$. D'autre part :

$$\int_a^b x^p \omega_n(x)\, \mu(x)dx = \sum_{j=0}^{n} A_j^{(n)}\, x_j^p\, \omega_n(x_j) = 0 \quad \text{car} \quad \omega_n(x_j) = 0 \,.$$

* Inversement supposons (i) et (ii) satisfaits. Soit $P \in \mathcal{P}_{2n+1}$

$$P = \omega_n\, Q + R \qquad \text{avec} \begin{cases} d^o\, R \leqslant n \\ d^o\, Q \leqslant n \end{cases} \qquad \text{(division euclidienne)}$$

donc $\displaystyle\int_a^b P(x)\mu(x)dx = \int_a^b Q(x)\omega_n(x)\mu(x)dx + \int_a^b R(x)\mu(x)dx$.

D'après (ii) $\displaystyle\int_a^b Q(x)\omega_n(x)\mu(x)dx = 0$.

Donc $\displaystyle\int_a^b P(x)\mu(x)dx = \int_a^b R(x)\mu(x)dx$.

D'après (ii) $\displaystyle\int_a^b R(x)\mu(x)dx = \sum_{j=0}^{n} A_j^{(n)}\, R(x_j)$.

Donc $\displaystyle\int_a^b P(x)\mu(x)dx = \sum_{j=0}^{n} A_j^{(n)}\, R(x_j)$.

Il reste à chercher si il existe des polynômes ω_n vérifiant (i) et (ii).

Proposition 5.7. Si $\mu(x) > 0$, il existe un polynôme $\omega_n(x) = (x-x_o)\ldots(x-x_n)$ tel que (i) et (ii) soient vérifiés et ce polynôme est unique.

Démonstration. * Unicité

Soient ω_n et ω_n^* deux polynômes répondant à la question. Posons

$\omega_n - \omega_n^* = S$. On a $d^o\ S \leqslant n$, donc d'après (ii) :

$$\int_a^b x^p\ S(x)\mu(x)dx = 0 \qquad \text{avec} \quad p \leqslant n \ .$$

Posons $S(x) = \sum_{p=o}^{n} \alpha_p\ x^p$. Alors

$$\sum_{p=o}^{n} \alpha_p \int_a^b x^p\ S(x)\mu(x)dx = \int_a^b (\sum_{p=o}^{n} \alpha_p\ x^p)S(x)\mu(x)dx = \int_a^b [S(x)]^2\mu(x)dx \ .$$

Mais $\sum_{p=o}^{n} \alpha_p \int_a^b x^p\ S(x)\mu(x)dx = 0$ car $p \in n$. Donc

$$\int_a^b [S(x)]^2\mu(x)dx = 0$$

ce qui implique $S \equiv 0$ car $\mu > 0$.

* Existence

Soit $\omega_n \in \mathcal{P}_{2n+1}$; $\int_a^b x^p\ \omega_n(x)\mu(x)dx = 0$ par hypothèse.

Définissons sur $\mathcal{P}_{2n+1}$ un produit scalaire de poids μ par :

$$(f,g)_\mu = \int_a^b f(x)g(x)\mu(x)dx \ ,$$

alors $$(f,f)_\mu = \int_a^b f^2(x)\mu(x)dx \ .$$

Par conséquent $\int_a^b x^p\ \omega_n(x)\mu(x)dx = 0 \iff (\omega_n,\ x^p)_\mu = 0 \iff \omega_n \perp x^p$, $p \leqslant n$.

La direction de ω_n est donc déterminée par le fait que ω_n est orthogonal à l'hyperplan engendré par $\{x^o, x^1, \ldots, x^n\}$. D'autre part, par construction le coefficient du terme de degré n dans ω_n est 1 . Cependant les racines $x_o, \ldots, x_n$ de ω_n pourraient être complexes. Il n'en est rien :

Lemme. ω_n a $(n+1)$ racine réelles dans $]a,b[$.

Soient $\xi_1, \ldots, \xi_m$ les racines de ω_n appartenant à $]a,b[$, de multiplicités impaires ; on suppose que ces racines sont distinctes.

On a $m \geqslant n+1$ et supposons $m \leqslant n$.

Soit $(x-\xi_1)\ldots(x-\xi_m)$ polynôme dont le degré est inférieur ou égal à n .

Donc : $$\int_a^b (x-\xi_1)\ldots(x-\xi_m)\ \omega_n(x)dx = 0\ .$$

$(x-\xi_1)\ldots(x-\xi_m)\ \omega_n(x)$ s'annule donc en $\xi_1 \ldots \xi_m$ et est de signe constant sur $]a,b[$. Il est donc identiquement nul ce qui est contradictoire avec l'hypothèse. Il y a donc exactement $(n+1)$ racines de ω_n appartenant à $]a,b[$ et ce sont les seules racines de ω_n .

Remarque. Les polynômes $\omega_o,\ldots,\omega_n$ forment les polynômes orthogonaux pour le poids μ et les points $x_o,\ldots,x_n$ sont les racines du polynôme orthogonal aux monômes $1, x,\ldots, x^n$.

(3) La 3e méthode qui va être décrite, utilise l'accélération de convergence. (Méthode de Romberg).

Il s'agit toujours de calculer $T = \int_a^b f(x)\mu(x)dx$ et le problème est le suivant : ayant calculé des valeurs approchées $T_o^o,\ldots, T_o^k$ de T on veut chercher à exploiter ces $(k+1)$ calculs approchés pour obtenir une approximation de T encore meilleure.

Une façon de résoudre ce problème est d'utiliser la technique d'accélération de convergence, présentée dans un cadre général.

* Partons d'une suite (T_o^k), (les T_o^k étant des valeurs approchées de T)

qui converge vers T .

Construisons le tableau triangulaire suivant, selon la formule :

$$\begin{array}{ccc} T_{m-1}^{k} & \searrow & \\ T_{m-1}^{k+1} & \longrightarrow & T_m^k \end{array} \qquad \boxed{T_m^k = \frac{4^m\, T^{k+1} - T_{m-1}^{k}}{4^m - 1}}$$

$$\begin{array}{cccccc}
T_o^o & & & & & \\
T_o^1 & T_1^o & & & & \\
T_o^2 & T_1^1 & T_2^o & & & \\
\vdots & \vdots & & \ddots & & \\
\vdots & \vdots & & & T_{k-1}^o & \\
T_o^k & T_1^{k-1} & \cdots\cdots & & T_{k-1}^1 & T_k^o \\
\vdots & & \cdots\cdots & & \cdots\cdots & \cdots\cdots \quad \ddots
\end{array}$$

On va montrer que si la suite (T_o^k) converge, il en est de même des autres suites "colonnes" (T_i^k) [i indice de la colonne] et également des suites constituées par les termes d'une même diagonale.

* Ce résultat, non limité aux quadratures est tout à fait général. Sous certaines conditions que nous ne développerons pas ici, les suites ainsi construites peuvent converger vers T beaucoup plus rapidement que la suite initiale (T_o^k). Il en est par exemple ainsi, pour une quadrature, lorsque T_o^k est la valeur obtenue par une formule des trapèzes à $2^k + 1$ points équidistants.

Cherchons tout d'abord une relation entre T_m^k , T_o^k, T_o^{k-m} .

On vérifie aisément sur le schéma triangulaire que T_m^k est une combinaison linéaire de $T_o^k, \dots, T_o^{k+m}$ à coefficients indépendants de k. Posons donc :

$$T_m^k = \sum_{i=o}^{m} C_{m,\, m-i} \, T_o^{k+i} \quad .$$

Changeons d'indice dans cette formule, on obtient :

d'une part $\qquad T_{m-1}^{k+1} = \sum_{i=o}^{m-1} C_{m-1,m-1-i} \, T_o^{k+1+i}$

d'autre part $\qquad T_{m-1}^{k} = \sum_{i=o}^{m-1} C_{m-1,m-1-i} \, T_o^{k+i}$

c'est-à-dire les 2 éléments à partir desquels on obtient T_m^k .

Ecrivons alors que $\quad T_m^k = \dfrac{4^m \, T_{m-1}^{k+1} - T_{m-1}^k}{4^m - 1} \iff (4^m-1)T_m^k = 4^m \, T_{m-1}^{k+1} - T_{m-1}^k$

On a : $(4^m-1) \sum_{i=o}^{m} C_{m,m-i} \, T_o^{k+i} = 4^m \sum_{i=o}^{m-1} C_{m-1,m-1-i} \, T_o^{k+1+i} - \sum_{i=o}^{m-1} C_{m-1,m-1-i} \, T_o^{k+i}$,

d'où si $m \geqslant 0$, $i = 0, \dots, m$, la relation de récurrence suivante entre les

coefficients $C_{\alpha,\beta}$:

$$(4^m - 1)C_{m,m-i} = 4^m \, C_{m-1,m-i} - C_{m-1,m-1-i} \quad .$$

Pour englober les cas limites, on pose : $C_{m-1,0} = C_{m-1,-1} = 0$.

Formons maintenant le polynôme associé aux $C_{m,k}$:

$$\Pi_m(X) = \sum_{k=o}^{m} C_{m,k} \, X^k \, .$$

On a la :

<u>Proposition</u> 6.1. Le polynôme $\Pi_m(X) = \sum_{k=o}^{m} C_{m,k} \, X^k$ vérifie la relation de récurrence :

$$\Pi_m(X) = \frac{4^m - X}{4^m - 1} \, \Pi_{m-1}(X)$$

et on a :

$$\Pi_m(X) = \frac{(1-\frac{X}{4^m})\dots(1-\frac{X}{4})}{(1-\frac{1}{4^m})\dots(1-\frac{1}{4})} \, \Pi_o(X)$$

avec $\qquad \Pi_o(X) = C_{oo} = 1$.

Considérons le polynôme $\Pi_{m-1}(X) = \sum_{k=o}^{m-1} c_{m-1,k} X^k$. On a :

$$(4^m - X)\Pi_{m-1}(X) = 4^m \sum_{k=o}^{m-1} c_{m-1,k} X^k - \sum_{k=o}^{m-1} c_{m-1,k} X^{k+1}$$

$$= 4^m \sum_{j=o}^{m-1} c_{m-1,m-1-j} X^{m-1-j} - \sum_{j=o}^{m-1} c_{m-1,m-1-j} X^{m-j}$$

(on a fait le changement d'indice $j = m-1-k$).

Le coefficient de X^{m-j} dans $(4^m - X)\Pi_{m-1}(X)$ soit $\dfrac{4^m c_{m-1,m-j} - c_{m-1,m-1-j}}{4^m - 1} = c_{m,m-j}$

est donc égal au coefficient de X^{m-j} dans $\Pi_m(X)$.

On a donc :

$$\Pi_m(X) = \frac{(1 - \frac{X}{4^m})\Pi_{m-1}(X)}{(1 - \frac{1}{4^m})} = \frac{(4^m - X)}{(4^m - 1)} \Pi_{m-1}(X) .$$

En "remontant" cette formule de récurrence, on obtient finalement en posant

$\Pi_o(X) = c_{oo} = 1$

$$\Pi_m(X) = \frac{(1 - \frac{X}{4^m})\dots(1 - \frac{X}{4})}{(1 - \frac{1}{4^m})\dots(1 - \frac{1}{4})} \Pi_o(X) .$$

Explicitons maintenant le coefficient $c_{m,k}$ de X^k dans $\Pi_m(X)$. On a le résultat suivant

<u>Proposition</u> 6.2. <u>Propriétés des coefficients</u> $c_{m,k}$

(i) $\sum_{k=o}^{m} c_{m,k} = \Pi_m(1) = 1$

(ii) $(-1)^k c_{m,k} > 0$

(iii) $\sum_{k=o}^{m} |c_{m,k}| = \sum_{k=o}^{m} (-1)^k c_{m,k} = \Pi_m(-1)$

$$= \frac{(1 + \frac{1}{4})\dots(1 + \frac{1}{4^m})}{(1 - \frac{1}{4})\dots(1 - \frac{1}{4^m})} \leqslant \frac{\prod_{\ell=o}^{\infty} (1 + \frac{1}{4^\ell})}{\prod_{\ell=o}^{\infty} (1 - \frac{1}{4^\ell})} \leqslant c_*$$

(On montre que $c_* < 2$).

(iv) $C_{m,k} = \dfrac{C_{m-1,k-1}}{1-4^k} = \left\{\dfrac{1}{(1-4)^k\ldots(1-4)}\right\} C_{m-k,o}$.

Démonstration. Les 3 premières propriétés sont immédiates. Montrons la 4^e.

On a $$\Pi_m(4z) = \frac{(1-\frac{z}{4^{m+1}})(1-\frac{z}{4^{m-2}})\ldots(1-\frac{z}{4})(1-z)}{(1-\frac{1}{4^m})\ldots\qquad\qquad(1-\frac{1}{4})}$$

Remarquons que : $1 - z = 1 - \frac{z}{4^m} + \frac{z}{4^m} - z$. Donc :

$$\Pi_m(4z) = \Pi_m(z) + \frac{(1-\frac{z}{4^{m-1}})\ldots(1-\frac{z}{4})(\frac{1}{4^m}-1)}{(1-\frac{1}{4^m})\ldots(1-\frac{1}{4})}\, z$$

$$= \Pi_m(z) - z\,\Pi_{m-1}(z)$$

d'où $$\sum_{k=o}^{m} C_{m,k}\, 4^k z^k = \sum_{k=o}^{m} C_{m,k}\, z^k - \sum_{k=o}^{m-1} C_{m-1,k}\, z^{k+1}$$

donc $(4^k - 1)C_{m,k} = -C_{m-1,k-1}$. Il en résulte que :

$$C_{m,k} = \frac{C_{m-1,k-1}}{1-4^k} = \frac{C_{m-2,k-2}}{(1-4^k)(1-4^{k-1})} = \left\{\frac{1}{(1-4^k)\ldots(1-4)}\right\} C_{m-k,o}$$

en remontant la récurrence.

Etudions maintenant la convergence de $\Pi_m(z)$ pour $z \in \mathbb{C}$.

Proposition 6.3. $\Pi_m(z)$ converge uniformément sur tout compact de $\mathbb{C}$ vers $\Pi(z)$.

Démonstration. $\mathrm{Log}\left|1 - \frac{z}{4^{m'}}\right| \leqslant \frac{|z|}{4^{m'}}$. Par conséquent $\prod_{m'=o}^{\infty} (1 - \frac{z}{4^{m'}})$ est un produit uniformément et normalement convergeant sur tout compact de $\mathbb{C}$. Il en est donc de même pour :

$$\frac{\prod_{m'=o}^{\infty}(1-\frac{z}{4^{m'}})}{\prod_{m'=o}^{\infty}(1-\frac{1}{4^{m'}})}\quad \Pi_o(X) = \lim_{m\to\infty} \Pi_m(z) .$$

Remarque. Comme les fonctions Π_m sont analytiques dans le plan, la fonction Π est aussi analytique et les suites dérivées k^e de Π_m, soient $\Pi_m^{(k)}(z)$, convergent uniformément vers la dérivée correspondante $\Pi^{(k)}(z)$ sur tout compact de $\mathbb{C}$. Résumons, on a :

Corollaire 6.3. $\Pi(z) = \lim\limits_{\substack{m\to\infty \\ z\in K\subset\mathbb{C}}} \Pi_m(z)$ est analytique. K compact de $\mathbb{C}$

Les $\Pi_m^{(k)}$ convergent uniformément vers $\Pi^{(k)}$ sur tout compact de $\mathbb{C}$.

Proposition 6.4. Quand $m \to \infty$, $C_{m,i} \to C_i$.

Démonstration. On a vu que $\Pi_m(z) = \sum\limits_{i=0}^{m} C_{m,i} z^i$ converge uniformément sur tout compact vers $\Pi(z) = \sum\limits_{i=0}^{\infty} C_i z^i$. Fixons i et prenons $m \geqslant i$. Alors :

$$\lim_{m\to\infty} C_{m,i} = \lim_{m\to\infty} \frac{1}{i!} \frac{d^i}{dz^i} \Pi_m(0)$$

soit
$$C_i = \frac{1}{i!} \frac{d^i}{dz^i} \Pi(0)$$

Corollaire 6.4.1. Quand $m \to \infty$, $C_{m,m-k} \to 0$.

Corollaire 6.4.2. Quand $m \to \infty$, $C_{m,k} \to C_k = C_o \left\{\frac{1}{(1-4^k)\ldots(1-4)}\right\}$.

Démonstration. C'est une conséquence immédiate du corollaire 6.4.1. et de (iv) dans la proposition 6.2.

Corollaire 6.4.3. $\Pi(z) = \sum\limits_{k=0}^{\infty} C_k z^k = C_o \sum\limits_{k=0}^{\infty} \frac{z^k}{(1-4)^k\ldots(1-4)}$

et
$$\Pi(1) = \lim_{m\to\infty} \Pi_m(1) = 1 .$$

Démonstration. Conséquence directe du corollaire 6.4.2.

Nous allons maintenant montrer deux propositions sur la convergence des T_m^k.

Proposition 6.5. (Convergence des colonnes)

Si la limite lorsque $k \to \infty$ de T_o^k existe (soit T cette limite) alors :

$\forall m \quad \lim\limits_{k\to\infty} T_m^k = T$.

<u>Démonstration</u>. On a vu que :

$$T_m^k = \sum_{i=0}^{m} C_{m,m-i}\, T_o^{k+i} \qquad , \quad T = \sum_{i=0}^{m} C_{m,m-i}\, T$$

Formons $\quad T_m^k - T = \sum_{i=0}^{m} C_{m,m-i}\, (T_o^{k+1} - T)$.

D'après les propositions précédentes (analycité et $C_{m,m-i} \to 0$)

$$\lim_{n\to\infty} T_m^k - T = 0$$

<u>Proposition</u> 6.6. (<u>Convergence des termes diagonaux</u>)

Si la limite de T_o^k existe et égale T quand $k \to \infty$, alors, pour tout

k fixé : $\lim\limits_{m\to\infty} T_m^k = T$.

<u>Démonstration</u>. Formons $\quad T_m^k - T = \sum_{i=0}^{m} C_{m,m-i}\, (T_o^{k+i} - T)$

$$= \sum_{i=0}^{m} C_{m,j}\, (T_o^{k+m-j} - T)$$

(on a posé $m - i = j$)

$\forall\, \varepsilon > 0$, $\exists\, i_\varepsilon$ t.q. $i \geqslant i_\varepsilon \implies |T_o^i - T| \leqslant \varepsilon$ par hypothèse

ou $\forall\, \varepsilon > 0$, $\exists\, i_\varepsilon$ t.q. $i \geqslant i_\varepsilon \implies |T_o^{k+i} - T| \leqslant \varepsilon$.

Pour $m \geqslant i_\varepsilon$, découpons la somme $\sum\limits_{i=0}^{m}$ en deux :

$$T_m^k - T = \sum_{i=0}^{i_\varepsilon} C_{m,m-i}\, (T_o^{k+i} - T) + \sum_{i=i_2}^{m} C_{m,m-i}\, (T_o^{k+i} - T) .$$

Donc $$|T_m^k - T| \leqslant \Big|\sum_{i=0}^{i_\varepsilon} C_{m,m-i}(T_o^{k+i} - T)\Big| + \sum_{i=i_2}^{m} |C_{m,m-i}|\, \varepsilon$$

$$\leqslant \Big|\sum_{i=0}^{i_\varepsilon} C_{m,m-i}(T_o^{k+i} - T)\Big| + C^*\, \varepsilon$$

Mais $\lim\limits_{m\to\infty} C_{m,m-i} = 0 \quad \forall i$. Faisons tendre m vers l'infini dans la dernière inégalité :

$$\limsup_{m\to\infty} \left|T_m^k - T\right| \leqslant \varepsilon\, C_* \,.$$

Comme $\varepsilon > 0$ est arbitraire, on a :

$$0 \leqslant \liminf_{m\to\infty} \left|T_m^k - T\right| \leqslant \limsup_{m\to\infty} \left|T_m^k - T\right| = 0$$

d'où $\lim\limits_{m\to\infty} T_m^k = T$.

CHAPITRE II

INTEGRATION NUMERIQUE DES EQUATIONS DIFFERENTIELLES

A THEOREMES D'EXISTENCES

Problème de Cauchy des équations différentielles.

α) Cas de $\mathbb{R}^N$.

Soit $\Omega \subset \mathbb{R} \times \mathbb{R}^N$ et $f : \Omega \to \mathbb{R}^N$.

On cherche un intervalle ouvert I et une fonction $y : I \subset \mathbb{R} \to \mathbb{R}^N$ vérifiant :

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

(ii) $y(t_o) = y_o$.

Plus précisément supposons f continue de Ω dans $\mathbb{R}^N$. Soit $(t_o, y_o) \in \Omega$. Peut-on trouver un intervalle ouvert I contenant t_o tel qu'il existe y appartenant à $\mathcal{C}^1(I, \mathbb{R}^N)$ vérifiant :

(i) $y'(t) = f(t, y(t)) \qquad \forall (t, y(t)) \in \Omega$, $t \in I$

(ii) $y(t_o) = y_o$.

Ce problème est appelé le problème aux valeurs initiales ou problème de Cauchy pour l'équation différentielle $y'(t) = f(t, y(t))$.

β) Cas d'un espace de Banach.

Plus généralement, soit E un espace de Banach et soit la fonction continue $f : \Omega \to E$ où Ω est un ouvert de $\mathbb{R} \times E$.

Peut-on trouver un ouvert I contenant t_o tel qu'il existe y appartenant à $\mathcal{C}^1(I, E)$ vérifiant :

(i) $y'(t) = f(t, y(t)) \qquad \forall (t, y(t)) \in \Omega$, $t \in I$

(ii) $y(t_o) = y_o$.

Si y existe on dit que y est solution sur I du problème de Cauchy.

Suivant l'hypothèse de départ, le problème de Cauchy admet une réponse plus ou moins précise. Nous allons nous placer dans les 2 cas suivants :

Hypothèses 1 $\quad H_1 \left\{ \begin{array}{l} E \text{ est de dimension finie} \\ f \text{ est continue sur } \Omega \end{array} \right.$

Hypothèses 2 $\quad H_2 \left\{ \begin{array}{l} E \text{ de dimension finie ou infinie} \\ f \text{ est continue sur } \Omega \\ f \text{ est localement lipschitzienne en } y \text{ sur } \Omega \end{array} \right.$

Définition 1. Fonction lipschitzienne.

On dit qu'une fonction $f : \Omega \to E$ continue est lipschitzienne en y pour la constante k si :

$$\begin{array}{l} \forall (t,y) \in \Omega \\ \forall (t,y_*) \in \Omega \end{array} \qquad \|f(t,y) - f(t,y_*)\| \leqslant k\|y-y^*\| .$$

Définition 2. Fonction localement lipschitzienne.

On dit qu'une fonction $f : \Omega \to E$ continue est localement lipschitzienne en y si pour tout point de Ω il existe un voisinage v de ce point dans lequel f est lipschitzienne en y.

Remarquons que :

Proposition. Si $f : \Omega \to E$ est continument différentiable en y, f est localement lipschitzienne en y dans Ω.

Démonstration. Soit $(t_o,y_o) \in \Omega$. Par hypothèse la différentielle f'_y existe et est continue. Donc il existe un voisinage fermé $\mathcal{V}$ de (t_o,y_o), contenu dans Ω, dans lequel :

$$\|f'_y\| \leqslant K .$$

Donc d'après le théorème des accroissements finis :

$$\begin{array}{l} \forall (t,y) \in \mathcal{V} \\ \forall (t,y_*) \in \mathcal{V} \end{array} \quad \begin{aligned} \|f(t,y_*) - f(t,y)\| &\leqslant \|f'_y\| \, \|y-y_*\| \\ &\leqslant K \, \|y-y_*\| \end{aligned}$$

On va maintenant montrer 2 résultats correspondant respectivement aux hypothèses H_1 et H_2.

<u>Théorème</u> 1. (Théorème d'existence ou théorème de Cauchy-Arzela).

Sous l'hypothèse H_1 le problème de Cauchy possède au moins une solution y définie sur un intervalle I contenant t_o.

<u>Démonstration</u>. On va seulement démontrer l'existence sur un intervalle $[t_o, t_o+h]$, le résultat complet étant facile à obtenir.

L'idée est de construire une solution approchée, puis de passer à la limite. La solution approchée sera constituée par une suite de fonctions continues, linéaires par morceaux sur $[t_o, t_o+h]$.

Commençons par déterminer un voisinage commode de $(t_o, y_o) = (t_o, y(t_o))$ dans lequel, on le verra, la solution est définie. La fonction f étant continue, il existe un voisinage $\mathcal{V}$ fermé de (t_o, y_o) tel que :

$$\|f(t,y)\| \leqslant M \qquad (t,y) \in \mathcal{V} .$$

Choisissons ensuite $h > 0$ assez petit de sorte que :

$$\mathcal{V}_h \subset \mathcal{V} \text{ avec } \mathcal{V}_h = \{(t,y) \mid |t-t_o| \leqslant h \ , \ \|y-y_o\| \leqslant M\,h\} .$$

On a donc : $\quad \|f(t,y)\| \leqslant M \qquad \forall (t,y) \in \mathcal{V}_h$.

Subdivisons $[t_o, t_o+h]$ en N intervalles de largeur $\frac{h}{N}$, h étant fixé :

$$t_o < t_1 < \ldots < t_N = t_o+h \ , \quad t_i = t_o + \frac{i}{N}\,h .$$

Considérons alors une solution approchée, continue, linéaire sur chaque intervalle $[t_i, t_{i+1}]$ définie de la façon suivant :

$$y_N(t_o) = y_o$$

$$y'_N(t) = \frac{y_N(t_{i+1}) - y_N(t_i)}{t_{i+1} - t_i} \qquad t \in]\, t_i, t_{i+1}[\, .$$

Une telle définition a un sens, si l'on montre que l'on peut définir y_N de proche en proche sur chaque intervalle $[t_i, t_{i+1}]$, c'est-à-dire si l'on montre que :

$$(t_j, y_N(t_j)) \in \mathcal{V}_h \, , \; 0 \leqslant j \leqslant i \implies (t, y_N(t)) \in \mathcal{V}_h \, , \; t \in [t_i, t_{i+1}] \, .$$

Supposons que : $(t_j, y_N(t_j)) \in \mathcal{V}_h$, $0 \leqslant j \leqslant i$.

On a alors : $\|y_N(t_{j+1}) - y_N(t_j)\| = \|f(t_j, y_N(t_j))\| \, . \, |t_{j+1}-t_j|$

$$\leqslant M|t_{j+1}-t_j| = M \frac{h}{N}$$

et ceci pour tout j tel que : $0 \leqslant j \leqslant i$.

Donc d'après l'hypothèse :

$$\|y_N(t_{i+1})-y_o\| \leqslant \sum_{j=o}^{i} \|y_N(t_{j+1})-y_N(t_j)\| \leqslant \sum_{j=o}^{i} M|t_{j+1}-t_i| = M\, N \frac{h}{N} = M\, h \, .$$

Comme $|t_{i+1}-t_o| \leqslant h$ par construction, on en déduit que :

$$(t_{i+1}, y_N(t_i)) \in \mathcal{V}_h$$

La définition de y_N a donc bien un sens.

Nous allons maintenant montrer que la suite $\{y_N\}$ ainsi définie tend uniformément sur $[t_o, t_o+h]$ vers une fonction y . Pour cela nous allons utiliser le théorème d'Ascoli.

Rappel. Théorème d'Ascoli.

On suppose que B est un espace de Banach et E un espace métrique compact. Pour qu'une partie H de l'espace de Banach $\mathcal{C}^0(E, B)$ soit relativement compacte il faut et il suffit que H soit équicontinue et que pour tout $x \in E$ l'ensemble $H(x)$ de tous les $f(x)$ tels que $f \in H$ soit relativement compact dans B.

Dans le cas présent : $E = I = [t_o, t_o+h]$ qui est bien compact

$B = E$ qui est un espace de Banach par hypothèse.

* <u>Montrons tout d'abord que</u> $H = \{y_N, N \in \mathbb{N}\} \subset \mathcal{C}^0(I,E)$ <u>est équicontinu sur</u> $[t_o, t_o+h]$.

Par définition $y_N'(t)$ existe p.p et pour $t \in]t_i, t_{i+1}[$:

$$\|y_N'(t)\| = \|f(t_i, y_N(t_i))\| \leqslant M .$$

Soient t, t_* quelconques dans $[t_o, t_o+h]$, alors :

$$y_N(t) - y_N(t_*) = \int_{t_*} y_N'(s)ds \qquad (1)$$

$$(1) \Longrightarrow \quad \|y_N(t) - y_N(t_*)\| = \left\|\int_{t_*}^{t} y_N'(s)ds\right\| \leqslant \sup_{t \in]t, t_*[} \|y_N'(t)\| . |t-t_*|$$

ainsi $\quad \forall\, t \in [t_o, t_o+h], \ \|y_N(t) - y_N(t_*)\| \leqslant M|t-t_*|$

$\forall\, t_* \in [t_o, t_o+h]$.

* <u>Montrons que</u> $H(t) = \{y_N(t) | y_N \in H\}$ $\forall t \in [t_o, t_o+h]$ <u>est relativement compact</u>.

E étant de dimension finie, il suffit, pour prouver que $H(t)$ est relativement compact, de montrer que $H(t)$ est borné.

Soit t quelconque dans $[t_o, t_o+h]$. Pour tout $y_N \in H$, nous avons :

$$y_N(t) = \int_{t_o}^{t} y_N'(s)ds + y_o$$

donc :

$$\|y_N(t)\| \leqslant \int_{t_o}^{t} \|y_N'(s)\|ds + \|y_o\|$$

$$\leqslant \sup_{s \in]t_o, t[} \|y_N'(s)\| . |t-t_o| + \|y_o\|$$

$$\leqslant M\,h + \|y_o\| < \infty .$$

Ainsi : $\forall\, t \in [t_o, t_o+h]$ les éléments $y_N(t)$ de $H(t)$ sont bornés par une quantité finie. Donc $H(t)$ est borné pour tout t de $[t_o, t_o+h]$.

* Nous pouvons alors appliquer le théorème d'Ascoli et en déduire que $H = \{y_N \,;\, N \in \mathbb{N}\}$ est relativement compact. Autrement dit il existe une fonction y de $\mathscr{C}^o(I,E)$ tel que :

$$y_\nu \to y \quad \text{uniformément sur} \quad I = [t_o, t_o+h] \,,$$

où $\{y_\nu\}$ est une suite extraite de $\{y_N\}$. En effet rappelons que :

$H \subset \mathscr{C}^1(I,E)$ relativement compact $\Longleftrightarrow$ toute suite de point de H possède une valeur d'adhérence dans $\mathscr{C}^1(I,E)$.

Dans ce qui suit on notera, pour simplifier, $\{y_N\}$ la suite extraite $\{y_\nu\}$.

On sait que par hypothèse : $y_N(t_o) = y_o$

$$\|y_N(t) - y_o\| \leqslant M\,h \qquad \forall\, t \in [t_o, t_o+h] \,.$$

Par passage à la limite on en déduit donc :

$$y(t_o) = y_o$$

$$\|y(t) - y_o\| \leqslant M\,h \,, \quad \forall\, t \in [t_o, t_o+h]$$

ce qui est équivalent à : $(t, y(t)) \in \mathcal{V}_h \subset \Omega \,,\ \forall\, t \in [t_o, t_o+h]$.

La limite y de la suite $\{y_N\}$ ne "sort" donc pas du voisinage $\mathcal{V}_h$ dans lequel on a défini les fonctions y_N .

* Montrons enfin que y satisfait (i) et (ii) c'est-à-dire est solution du problème de Cauchy. On sait que les conditions (i) et (ii) sont équivalentes à :

$$y(t) = y_o + \int_{t_o}^{t} f(s, y(s))ds \quad , \quad \forall\, t \in [t_o, t_o+h] \,.$$

Or si $t \in\,]t_i, t_{i+1}[$, par construction :

$$y_N(t) - y_N(t_i) = (t-t_i)\, f(t_i, y_N(t_i))$$

$$y_N(t_i)-y_N(t_{i-1}) = (t-t_{i-1})\, f(t_{i-1}, y_N(t_{i-1}))$$

$$\cdots\cdots$$

$$y_N(t_1) - y_o = (t_1-t_o) f(t_o, y_N(t_o)) .$$

Sommons ces égalités, on obtient :

$$y_N(t) - y_o = (t-t_i)\, f(t_i, y_N(t_i)) + \sum_{j=1}^{i} (t_j-t_{j-1})\, f(t_{j-1}, y_N(t_{j-1})) .$$

Posons $I_{1,N} = (t-t_i)\, f(t_i, y_N(t_i)) + \sum_{j=1}^{i} (t_j-t_{j-1})\, f(t_{j-1}, y(t_{j-1}))$

$$I_{2,N} = (t-t_i)[f(t_i, y_N(t_i)) - f(t_i, y(t_i))] + \sum_{j=1}^{i} (t_j-t_{j-1})[f(t_{j-1}, y_N(t_{j-1})) - f(t_{j-1}, y(t_{j-1}))] .$$

On vérifie que : $y_N(t) = y_o + I_{1,N} + I_{2,N}$ (1) .

Il est clair que, f étant continue, $I_{1,N}$ est une somme de Riemann pour l'intégrale :

$$\int_{t_o}^{t} f(s, y(s))ds$$

et pour la subdivision $t_o < t_1 < \ldots < t_i < t$. Donc :

$$\lim_{N\to\infty} I_{1,N} = \int_{t_o}^{t} f(s, y(s))ds .$$

Par ailleurs, f étant continu et y_N convergeant uniformément vers y sur $[t_o, t_o+h]$ il en résulte que :

$$f(t, y_N(t)) \to f(t, y(t)) \text{ uniformément sur } [t_o, t_o+h] .$$

Autrement dit :

$$\forall \varepsilon > 0 , \exists N_\varepsilon \text{ t.q. } N \geqslant N_\varepsilon \Longrightarrow \|f(s, y_N(s)) - f(s, y(s))\| \leqslant \varepsilon , \forall s \in [t_o, t_o+h] .$$

Par conséquent, comme t_j pour $1 \leqslant j \leqslant i$ appartient à $[t_o, t_o+h]$ pour $N \geqslant N_\varepsilon$ on a :

$$\|I_{2,N}\| \leqslant |t-t_i| . \|f(t_i,y_N(t_i)) - f(t_i,y(t_i))\| + \sum_{j=1}^{i} |t_j-t_{j-1}| . \|f(t_{j-1},y_N(t_{j-1})) - f(t_{j-1},y(t_{j-1}))\|$$

d'où $\|I_{2,N}\| \leqslant \varepsilon\, [|t-t_i| + |t_i-t_{i-1}| + \ldots + |t_1-t_0|] = \varepsilon |t-t_0|$.

Ainsi : $\forall\ \varepsilon > 0$, $\exists\ N_\varepsilon$ t.q. $N \geqslant N_\varepsilon \Longrightarrow \|I_{2,N}\| \leqslant |t-t_0|\ \varepsilon$

par conséquent $\lim\limits_{N\to\infty} I_{2,N} = 0$.

Faisons alors tendre N vers l'infini dans (1). Par passage à la limite il en résulte que :

$$y(t) = y_0 + \int_{t_0}^{t} f(s,y(s))ds \ , \quad \forall\ t \in [t_0, t_0+h]$$

ce qui signifie que : $y \in \mathcal{C}^1(I,E)$ est solution du problème de Cauchy. Le théorème 1 est ainsi démontré.

<u>Théorème</u> 2. (Théorème d'existence et d'unicité ou théorème de Cauchy-Lipschitz).

Sous l'hypothèse (H_2), il existe un intervalle ouvert I contenant t_0 , tel que le problème de Cauchy possède une solution <u>unique</u> sur I .

On va utiliser la méthode des approximations successives et le théorème du point fixe.

<u>Rappel</u>. <u>Théorème du point fixe</u>.

Soit $\mathcal{B}$ une boule fermée d'un espace de Banach et $T : \mathcal{B} \to \mathcal{B}$, une contraction stricte $(\|Ty - Tz\| \leqslant C\|y-z\|$ avec $C < 1)$.

Alors il existe un y <u>unique</u> de $\mathcal{B}$ tel que : $Ty = y$.

Nous allons définir une application T vérifiant les hypothèses du théorème.

* Pour cela déterminons tout d'abord un voisinage $\mathcal{V}_h$ convenable. On a vu que $\mathcal{V}_h$ est déterminé par la donnée du nombre h . Il existe un voisinage $\mathcal{V}$

de (t_o, y_o) tel que :

$$\|f(t,y)\| \leqslant M \quad \text{pour} \quad (t,y) \in \mathcal{V} .$$

Donc d'après le théorème des accroissements finis, il existe un nombre

$k = \|f'_y\|$ tel que :

$$\|f(t,y) - f(t,y_*)\| \leqslant k\|y-y_*\| \qquad (t,y) \in v \ , \ (t,y_*) \in \mathcal{V}.$$

Définissons alors $\mathcal{V}_h$ comme :

$$\mathcal{V}_h = \{(t,y) \in \mathcal{V} | \ |t-t_o| \leqslant h \ , \ \|y-y_o\| \leqslant M\,h\} \qquad \text{avec} \quad h < \frac{1}{k} \ .$$

Posons d'autre part $I =]t_o-h, \ t_o+h[$.

* Définissons maintenant T . On sait que :

$$\left.\begin{array}{l} y \in \mathcal{C}^1(I,E) \\ \{(t,y(t)) | t \in I\} \subset \Omega \\ y(t) = y_o + \displaystyle\int_{t_o}^{t} f(s,y(s))ds \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{l} y \text{ est solution du problème} \\ \text{de Cauchy} \end{array}\right.$$

Posons donc : $T : y \mapsto Ty = z$ avec $z(t) = y_o + \displaystyle\int_{t_o}^{t} f(s,y(s))ds$.

T satisfait aux hypothèses du théorème du point fixe.

(1) $T : \mathcal{B} \to \mathcal{B}$ où $\mathcal{B} = \mathcal{C}_h(\overline{I},E) = \{y \in \mathcal{C}^1(\overline{I},E) | \ \|y(t)-y_o\| \leqslant M\,h \ , \ t \in \overline{I}\}$.

Soit $y \in \mathcal{C}_h(\overline{I},E)$. Appliquons T à y : $Ty = z$. $z \in \mathcal{C}^1(\overline{I},E)$ par construction.

De plus :

$$\|z(t) - y_o\| = \left\|\int_{t_o}^{t} f(s,y(s))ds\right\| \leqslant \int_{t_o}^{t} \|f(s,y(s))\| ds \leqslant M(t-t_o)$$

donc, comme $t \in [t_o-h, \ t_o+h]$: $\|z(t) - y_o\| \leqslant M\,h$.

Par conséquent $(s,y(s)) \in \mathcal{V}_h$ et $z \in \mathcal{C}_h(\overline{I},E)$.

(2) T est contractante.

Soient $Ty = z$, $Tx = v$ avec $(t,y) \in \mathcal{V}_h$, $(t,x) \in \mathcal{V}_h$, où $\mathcal{V}_h$ est le voisinage défini précédemment et où t appartient à $\overline{I}$.

Formons :

$$z(t) - v(t) = \int_{t_o}^{t} (f(s,y(s)) - f(s,x(s)))ds \quad .$$

On a alors :

$$\|z(t) - v(t)\| = \|\int_{t_o}^{t} (f(s,y(s)) - f(s,x(s)))ds\|$$

$$\leqslant \int_{t_o}^{t} \|f(s,y(s)) - f(s,x(s))\| \, ds \; .$$

Soit d'après l'inégalité des accroissements finis :

$$\|z(t) - v(t)\| \leqslant k \int_{t_o}^{t} \|y(s) - x(s)\| \, ds \qquad \text{(rappelons que } k = \|f'_y\|)$$

$$\leqslant k \; (t-t_o) \; \|y-x\|_{\mathcal{C}^1(\overline{I},E)}$$

d'où comme $t \in [t_o-h, \; t_o+h]$:

$$\|z(t) - v(t)\| < k \; h \; \|y-x\|_{\mathcal{C}^1(\overline{I},E)}$$

Mais cette inégalité est vraie pour tout t de $[t_o-h, \; t_o+h]$.

On peut donc écrire :

$$\|z - v\|_{\mathcal{C}^1(\overline{I},E)} \leqslant k \; h \; \|y - x\|_{\mathcal{C}^1(\overline{I},E)}$$

Soit encore par définition de z et v :

$$\|Ty - Tx\|_{\mathcal{C}^1(\overline{I},E)} \leqslant k \; h \; \|y - x\|_{\mathcal{C}^1(\overline{I},E)}$$

Mais par hypothèse (construction de $\mathcal{V}_h$) $k \; h < 1$; T est alors bien contractante.

On peut alors appliquer le théorème du point fixe et conclure qu'il existe un point y <u>unique</u> de $\mathcal{B} = \mathcal{C}_h(\overline{I},E)$ tel que $Ty = y$, autrement dit que :

$$y \in \mathcal{C}^1(\overline{I},E) \;\; , \;\; \{(t, \; y(t)) | \; t \in \overline{I}\} \subset \mathcal{V}_h \subset \Omega$$

et

$$y(t) = y_o + \int_{t_o}^{t} f(s,y(s)) \; ds$$

par définition de T .

<u>Corollaire</u> 2 . Supposons les hypothèses H_2 réalisées :

soit y (resp. z) la solution du problème de Cauchy sur l'intervalle ouvert J contenant t_o (resp. l'intervalle ouvert K contenant t_o).

Alors $y = z$ sur $J \cap K$.

<u>Démonstration</u>. Considérons $\mathfrak{U} = \{t \in J \cap K \mid y(t) = z(t)\}$.

$$\forall t_1 \in J , \exists h \quad t.q. \quad \begin{cases} y' = f(t,v) \\ v(t_n) = y(t_1) \end{cases} \quad \text{a une solution unique}$$

sur $]t_1-h, t_1+h[$ d'après le théorème de Cauchy-Lipschitz. On a donc $y = z$ sur $]t_1-h, t_1+h[$ et :

$$\mathfrak{U} \supset]t_1-h, t_1+h[$$

J est ouvert, $\mathfrak{U}$ est fermé et non vide, donc par connexité $\mathfrak{U} = J \cap K$.

Les théorèmes de Cauchy-Arzela et Cauchy-Lipschitz donnent une solution définie sur un petit intervalle $[t_o, t_o+h]$ ou $]t_o-h, t_o+h[$.

Le problème qui se pose est d'examiner si on peut prolonger les solutions, dont on a vu l'existence (et l'unicité) au delà de ces petits intervalles et si oui sous quelles conditions.

B PROLONGEMENT DES SOLUTIONS

a) <u>Solutions maximales</u>.

Nous allons définir une relation d'ordre simple sur l'ensemble des solutions du problème de Cauchy.

On note $y < z$ lorsque : $y : I \to E$ et $z : J \to E$ solution du problème de Cauchy respectivement définies sur I et J sont telles que $I \subset J$.

Cette relation d'ordre est en fait la relation d'inclusion sur les graphes des solutions.

Définition 2.1. On appelle solution maximale, tout élément maximal de l'ensemble des solutions ainsi ordonné.

Lemme 2.2. L'ensemble des solutions ordonné par α est inductif.

Démonstration. Soit $\{y_\alpha, I_\alpha\}_{\alpha\in\Lambda}$ une famille totalement ordonnée de solutions.

On a donc : $\forall \alpha \in \Lambda$ soit $y_\alpha < y_\beta$ soit $y_\alpha > y_\beta$.

$\forall \beta \in \Lambda$

Soit $I = \bigcup_{\alpha\in\Lambda} I_\alpha$. I est un intervalle ouvert et $t_o \in I$.

Soit $y : I \to E$ t.q. $\forall t \in I$, $y(t) = y_{\alpha_o}(t)$ où α_o est un élément de Λ tel que $t \in I_\alpha$.

Ces définitions sont comparables car :

$$y \in \mathscr{C}^1(I,E),\ \{(t, y(t)) \mid t \in I\} \subset \Omega$$

$$(y_\alpha, I_\alpha) < (y,I) \qquad \forall \alpha .$$

Donc y est solution sur I du problème de Cauchy et la famille $\{y_\alpha, I_\alpha\}_{\alpha\in\Lambda}$ admet une borne supérieure (y, I).

Proposition 2.3. Sous les hypothèses H_1 ou H_2 le problème de Cauchy admet des solutions maximales.

Sous l'hypothèse H_2 il existe une solution maximale unique.

Démonstration. L'existence résulte du Lemme de Zorn et du Lemme 2.2.. Il reste à démontrer l'uncitié dans le cas de l'hypothèse H_2.

Soient (y,I) et (z,J) deux solutions maximales. D'après le corollaire 2, on a $y = z$ sur $J \cap I = K$. Supposons que $I \neq J$. Soit w la solution du problème de Cauchy sur $K = I \cap J$. Alors :

$$(w,K) \underset{\neq}{>} (y,I) \quad \text{et} \quad (w,K) \underset{\neq}{>} (z,J)$$

ce qui est contradictoire avec le fait que (y,I) et (z,J) soient des solutions maximales. Donc $I = J$ et $y = z$ partout.

Résumons le théorème de Cauchy-Lipschitz et la proposition 2.3., on a le théorème fondamental suivant :

<u>Théorème</u> 2.4. Sous l'hypothèse H_2 le problème de Cauchy admet une solution maximale <u>unique</u>.

b) <u>Critère de prolongement</u>.

<u>Théorème</u> 2.5. Sous l'une des hypothèses H_1 ou H_2, soit y une solution du problème de Cauchy sur $]a,b[$. Alors y est prolongeable au delà de b (resp. en deçà de a) si et seulement si l'une des valeurs d'adhérence de $\{t,y(t)\}$ quand $t \to b-o$ (resp. $t \to a+o$) est dans Ω.

Autrement dit, le seul cas où le prolongement n'est pas possible est celui où toutes les valeurs d'adhérence sont sur $\partial\Omega$.

<u>Corollaire</u> 2.6. Sous l'une des hypothèses H_1 ou H_2 et si $\Omega = \mathbb{R} \times E$ la solution y du problème de Cauchy sur $]a,b[$ est prolongeable au delà de b (resp. en deçà de a) si et seulement si :

$$\underline{\lim}_{t \to b-o} \|y(t)\| < +\infty \quad (\text{resp. } \underline{\lim}_{t \to a+o} \|y(t)\| < +\infty) .$$

Autrement dit, le seul cas où y n'est pas prolongeable est celui où

$$\lim_{t\to b-o} \|y(t)\| = \infty \quad (\text{resp. } \lim_{t\to a+o} \|y(t)\| = \infty) .$$

<u>Démonstration</u>. (Théorème 2.5. et Corollaire 2.6.).

– Supposons y prolongeable.

Il existe donc $z : J \to E$, $J \supset]a,b[$ tel que :

$$y = z \text{ sur }]a,b[.$$

Pour $t \to b-o$, $y(t) = z(t) \to z(b)$ et donc $(b, z(b)) \in \Omega$.

– Inversement supposons que (b,β), valeur d'adhérence pour $t \to b-o$, soit dans Ω .

On va montrer essentiellement que : $\beta = \lim_{t\to b-o} y(t)$.

Soit $h > 0$. h définit le voisinage $\mathcal{V}_h$ tel que :

$$\mathcal{V}_h = \{(t, y(t)) \mid |t-b| \leqslant h , \|y-\beta\| \leqslant M h\} .$$

On sait que, f étant continue, il existe un voisinage $\mathcal{V}$ tel que :

$$\mathcal{V} = \{(t,\xi) \mid \|f(t,\xi)\| \leqslant M\} .$$

Or on a $\mathcal{V}_h \subset \mathcal{V}$ et par conséquent : $\forall (t,\xi) \in \mathcal{V}_h$ $\|f(t,\xi)\| \leqslant M$.

Définissons de plus : $\mathcal{W}_h = \mathcal{V}_{\frac{h}{3}} = \{(t,y(t)) \mid |t-b| \leqslant \frac{h}{3} , \|y-\beta\| \leqslant M \frac{h}{3}\}$.

$\mathcal{W}_h$ rencontre le graphe de y :

$$\exists\, b_1 \text{ t.q. } 0 \leqslant b-b_1 \leqslant \frac{h}{3} \quad \|y(b_1) - \beta\| \leqslant M \frac{h}{3} .$$

* Montrons que $\|y(t) - y(b_1)\| \leqslant \frac{2}{3} M h$ pour $t \in [b_1,b[$.

Pour $t = b_1$, c'est vérifié.

Cherchons alors l'intervalle maximum $[b_1, b_2]$ sur lequel l'inégalité a lieu.

Supposons $b_2 < b$, alors pour $t \in [b_1, b_2]$:

$$\|y(t)-\beta\| \leqslant \|y(t)-y(b_1)\| + \|y(b_1)-\beta\| \leqslant \frac{2}{3} M h + \frac{1}{3} M h = M h .$$

Or $0 \leqslant t-b_1 \leqslant b_1-b \leqslant M\frac{h}{3}$. Donc : $(t, y(t)) \in \mathcal{V}_h$ pour $t \in [b_1, b_2]$ et par conséquent : $y(b_2) - y(b_1) = \int_{b_1}^{b_2} f(s, y(s))ds$. Donc

$$\|y(b_2) - y(b_1)\| \leqslant (b_1-b_2)M \leqslant \frac{2}{3} M h .$$

Par continuité de y il en résulte :

$$\|y(t) - y(b_1\| \leqslant \frac{2}{3} M h \quad \text{pour } t \in [b_1, b_2+\delta[\ , \quad \delta > 0$$

ce qui est contradictoire avec la définition de b_2 . Donc $b_2 = b$ et on a alors :

$$\|y(t) - y(b_1)\| \leqslant \frac{2}{3} M h \quad \text{pour} \quad t \in [b_1, b[\ .$$

* Montrons que $y(t)$ est une suite de Cauchy pour $t \to b-o$. On a vu que pour $t \in [b_1, b[$:

$$\|y(t)-\beta\| \leqslant \|y(t) - y(b_1)\| + \|y(b_1)-\beta\| \leqslant M h$$

de sorte que : $(t, y(t)) \in \mathcal{V}_h$ pour $t \in [b_1, b[$.

Alors si $t \in [b_1, b[$, $t' \in [b_1, b[$:

$$\|y(t) - y(t')\| \leqslant \int_{t'}^{t} \|f(s, y(s))\|ds \leqslant |t-t'| M$$

$y(t)$ est donc bien une suite de Cauchy pour $t \to b-o$. y étant à valeurs dans E qui est complet par hypothèse, il en résulte que $y(t)$ a une limite quand $t \to b-o$. Ce ne peut être que β .

* Montrons enfin que y est prolongeable au delà de b .

Traduisons l'existence de solution au voisinage de b :

$$\exists\ \delta > 0 \ , \exists\ z \text{ définie sur } [b, b+\delta[\text{ t.q. } \begin{array}{l} z'(t) = f(t, z(t)) \\ z(b) \ = \beta \end{array} .$$

Définissons alors $\tilde{y}$ comme suit :

$$\tilde{y} = \begin{cases} y & \text{sur }]a,b[\\ z & \text{sur } [b, b+\delta[\ . \end{cases}$$

On a alors : $\tilde{y}(t) = y_o + \int_{t_o}^{t} f(s, y(s))ds$ pour $t \in]a, b+\delta[$, et donc $y \in \mathcal{C}^1(]a, b+\delta[, E)$ et $\tilde{y}$ est solution du problème de Cauchy sur $]a, b+\delta[$; $\tilde{y}$ est aussi un prolongement strict de y .

Cas particulier.

Supposons en plus de l'hypothèse H_1 que :

α) $\Omega = \mathbb{R} \times E$

β) f est lipschitzienne en y sur tout l'espace $\mathbb{R} \times E$:

$$\|f(t,y) - f(t,y_*)\| < L\|y - y_*\| \qquad \forall\ t \in \mathbb{R} ,\ \forall\ y, y_* \in E .$$

On a alors la :

Proposition 2.7. La solution du problème de Cauchy :

$$y'(t) = f(t, y(t)),\ \ y(t_o) = y_o \text{ est définie sur tout } \mathbb{R} .$$

Démonstration.

Montrons que si y est solution sur $[t_o, t_1]$ alors $y(t)$ est borné pour $t \to t_1$, ce qui entrainera que la solution est prolongeable au delà (et en deçà) de t_1 . On a :

$$y'(t) = f(t, y(t)) \implies \|y'(t)\| = \|f(t, y(t))\| \qquad (1)$$

f étant lipschitzienne en y :

$\|f(t,y) - f(t,o)\| \leqslant L\|y\|$ d'où

$\|f(t,y) \leqslant L\|y\| + \|f(t,o)\|$ soit d'après (1) :

$\|y'(t)\| \leqslant L\|y(t)\| + \|f(t,o)\|$.

Posons $z(t) = \|y'(t)\|$. On en déduit que :

$$z'(t) \leqslant \|f(t,o)\| + L\, z(t)$$

soit : $\quad z'(t) - L\, z(t) \leqslant \|f(t,o)\|$.

Multiplions les deux membres de cette inégalité par e^{-Lt}

$$e^{-Lt}(z'(t) - L\, z(t)) \leqslant \|f(t,o)\|\, e^{-Lt}$$

soit : $\quad \frac{d}{dt}[z(t)\, e^{-Lt}] \leqslant \|f(t,o)\|\, e^{-Lt}$

intégrons cette inégalité entre t_o et s :

$$z(s)\, e^{-Ls} - z(t_o)e^{-Lt_o} \leqslant \int_{t_o}^{s} \|f(t,o)\|\, e^{-Lt}\, dt ,$$

soit $\quad z(s)\, e^{-Ls} \leqslant \|y_o\|\, e^{-Lt_o} + \int_{t_o}^{t} \|f(t,o)\|\, e^{-Lt}\, dt$

car $\quad z(t_o) = \|y(t_o)\|$

$\quad\quad = \|y_o\|$.

Finalement, on a en multipliant les 2 membres par $e^{Ls} > 0$

$$z(s) \leqslant \|y_o\|\, e^{L(s-t_o)} + \int_{t_o}^{s} \|f(t,s)\|\, e^{L(s-t)} dt .$$

Faisons tendre s vers t_1-o . On a, à la limite :

$$\lim_{s\to t_1-o} z(s) = \|y(t_1-o)\| \leqslant \|y_o\|\, e^{L(t_1-t_o)} + \int_{t_o}^{t_1-o} \|f(t,s)\|\, e^{L(t_1-t_o)} dt$$

$$\leqslant \|y_o\|\, e^{L(t_1-t_o)} + M \int_{t_o}^{t_1-o} e^{L(t_1-t_o)}\, dt$$

quantité qui est finie donc $y(t)$ est borné quand $t \to t_1$.

C RESOLUTION NUMERIQUE (Méthodes de différence finie)

1 - Les méthodes à un et plusieurs pas.

1.1) Méthodes à un pas.

Nous nous plaçons dans le cas scalaire $(E = \mathbb{R}^N)$ du problème de Cauchy

(1,1) $\quad y'(t) = f(t, y(t))$

(1,2) $\quad y(t_o) = y_o$ donné .

On suppose qu'il y a existence et unicité de y dans $[a,b]$. On introduit un pas $h = \frac{b-a}{N}$ et l'on va d'abord "intégrer" (1,1) dans l'intervalle $(a, a+h)$ en remplaçant la courbe par sa tangente au point a, etc... Posons :

$$(1,3) \qquad \begin{cases} t_j = a + jh & 0 \leqslant j \leqslant N \\ t_o = a \quad t_N = b \end{cases}$$

Désignons par :

$$(1,4) \qquad y_n = \text{approximation de } y(t_n).$$

Alors la méthode précédente conduit au schéma suivant (d'Euler) :

$$(1,5) \qquad y_{n+1} = y_n + h\, f(t_n, y_n) \ , \qquad n \geqslant 0 \ ;$$

$y(t_o) = y_o$ étant connu, les équations (1,5) définissent de proche en proche $y_1, y_2, \ldots$; de façon générale y_{n+1} est déterminé à partir de la connaissance de y_n par (1,5) ; c'est ce que l'on appelle une méthode à <u>un pas</u>.

Il reste maintenant à savoir si y_n est vraiment une approximation de $y(t_n)$ et si $|y_n - y(t_n)| \to 0$ lorsque $h \to 0$, $t_N = a + Nh$ étant fixé.

Naturellement on peut espérer améliorer la précision par une modofication convenable de la méthode. C'est l'objet du n° 2.

1.2) <u>Méthodes à plusieurs pas</u>.

Intégrons l'équation (1,1) sur l'intervalle (t_{n-p}, t_{n+1}), p entier fixé (ceci n'a de sens que si $n \geqslant p$). Il vient :

$$(1,6) \qquad y(t_{n+1}) - y(t_{n-p}) = \int_{t_{n-p}}^{t_{n+1}} f(t, y(t))dt \ .$$

On peut ensuite calculer l'intégrale de droite en utilisant une formule de quadrature [cf. Chap. I § C] ; cela conduira par exemple à :

$$(1,7) \qquad y(t_{n+1}) - y(t_{n-p}) \approx \sum_{j=n-p}^{n+1} A_j\, f(t_j, y(t_j)) \ ,$$

ou $(1,8)$ $$y(t_{n+1}) - y(t_{n-p}) \approx \sum_{j=n-p}^{n+0} B_j\, f(t_j,\, y(t_j))\ .$$

L'équation (1,7) (resp. (1,8)) conduit aux <u>équations aux différences</u> :

$$(1,9) \qquad y_{n+1} - y_{n-p} = \sum_{j=n-p}^{n+1} A_j\, f(t_j,\, y_j)$$

resp.

$$(1,10) \qquad y_{n+1} - y_{n-p} = \sum_{j=n-p}^{n+0} B_j\, f(t_j,\, y_j))\ .$$

La détermination de y_{n+1} à l'aide de (1,9) ou (1,10) suppose connus $u_n, u_{n-1}, \ldots, u_{n+p}$, soit $p+1$ valeurs antérieures. C'est ce que l'on appelle <u>méthode à (p+1) pas</u>. Une différence essentielle apparait entre (1,9) et (1,10) ; la formule (1,10) donne <u>explicitement</u> y_{n+1} au moyen de $y_n, \ldots, y_{n-p}$: c'est une <u>méthode explicite</u>.

Par contre, la formule (1,9) ne donne y_{n+1} <u>qu'implicitement</u> ou après la résolution d'une équation (dont il faut montrer la possibilité). C'est une <u>méthode implicite</u>.

<u>Remarque</u> 1.1. Dans chacune des méthodes (1,9) (1,10) à (p+1) pas, apparait la difficulté du démarrage. Seul y_o étant donné on ne peut calculer y_1 au moyen de (1,9) ou (1,10) ; ces formules ne sont applicables qu'une fois $y_o, \ldots, y_p$ en notre possession. Il faut donc utiliser d'autres méthodes pour calculer $y_1, \ldots, y_p$.

<u>Remarque</u> 1.2. Tout ce qui vient d'être dit s'adapte sans difficulté au cas des systèmes d'équations différentielles.

2 - Méthodes à un pas.

2.1) Méthode d'Euler modifiée.

Reprenons (1,5) ; on peut l'obtenir comme suit : partant de la relation

$$y(t_{n+1}) - y(t_n) = \int_{t_n}^{t_{n+1}} f(t, y(t))dt.$$

On utilise la formule de la moyenne :

$$(2,1) \qquad y(t_{n+1}) - y(t_n) = h\, f(\bar{t}_n, y(\bar{t}_n)) , \quad t_n < \bar{t}_n < t_{n+1} .$$

La méthode d'Euler consiste à remplacer $\bar{t}_n$ par t_n d'où (1,5). Si on remplace $\bar{t}_n$ par $t_n + \frac{h}{2}$ et $y(t_n + \frac{h}{2})$ par sa valeur approchée calculée en (1,5) à savoir :

$$y_n + \tfrac{1}{2} h\, f(t_n, y_n)$$

il vient

$$(2,2) \qquad y_{n+1} - y_n = h\, f(t_n + \frac{h}{2} , y_n + \tfrac{1}{2} h\, f(t_n, y_n)) .$$

On ajoute à ce système (explicite) la condition initiale

$$(2,3) \qquad y_o \text{ connu.}$$

2.2) Méthode générale à un pas.

Soit $\Phi(t, y ; h)$ une fonction donnée pour $t \in [a,b]$, $y \in \mathbb{R}$, $h \geqslant 0$ et que nous supposerons toujours au moins continue en ces 3 variables. Alors les schémas (1,5) et (2,2) entrent tous deux dans le schéma général suivant :

$$(2,4) \qquad y_{n+1} = y_n + h\, \Phi(t_n, y_n ; h) .$$

Dans le cas d'Euler on a :

$$\Phi(t, y ; h) = f(t, y) \quad \text{(donc indépendant de } h!) ;$$

dans le cas d'Euler modifié (2,2) on a :

$$\Phi(t, y ; h) = f(t + \tfrac{1}{2} h, y + \frac{h}{2} f(t,y)) .$$

La condition minimum à supposer sur Φ est :

$$(2,7) \qquad \Phi(t, y ; 0) = f(t, y) .$$

En effet il faut imposer que $y_n \to y(t_n)$ et $\frac{y_{n+1} - y_n}{h} \to y'(t_n)$ lorsque $h \to 0$, $t_n = a + N h$ fixé ; d'après (2,4) et puisque $y'(t_n) = f(t_n, y(t_n))$ il faut donc que :

$$\Phi(t_n, y(t_n) ; 0) = f(t_n, y(t_n))$$

et ceci quelle que soit la donnée initiale y_o , d'où (2,7) .

Notre problème est maintenant le suivant : comment choisir $\Phi(t, y ; h)$. Nous allons examiner ce problème à partir de l'étude de la <u>convergence</u> et de <u>l'estimation de l'erreur</u>.

2.3) <u>Estimation de l'erreur</u>.

On introduit <u>l'erreur de discrétisation</u>

$$(2,8) \qquad e_n = y_n - y(t_n) .$$

Nous allons démontrer la

<u>Proposition</u> 2.1. On suppose que les conditions suivantes ont lieu :

(i) Φ satisfait à (2,7), $\Phi(t, y ; 0) = f(t, y)$

(ii) $\begin{cases} |\Phi(t, y ; h) - \Phi(t, y^* ; h)| \leqslant L_1 |y-y^*| \\ \text{pour } t \in [a,b], y, y^* \in \mathbb{R} \text{ (ou dans un compact convenable)} \\ \text{et pour } 0 \leqslant h \leqslant h_o \ (\leqslant b-a) ; \end{cases}$

(iii) $\begin{cases} |\Phi(t, y(t) ; h) - \frac{y(t+h) - y(t)}{h}| = \\ |\Phi(t, y(t) ; h) - \frac{1}{h}\int_t^{t+h} f(\xi, y(\xi))d\xi| \leqslant L_2 h^p \\ \text{pour toute solution } y \text{ régulière de } y' = f(t, y(t)). \end{cases}$

<u>Conclusion</u>. Il existe une constante K telle que

(2,9) $\quad |e_n| \leqslant K\,h^p$.

Démonstration.

1/ De (2,4) et de la formule : $y(t_{n+1}) - y(t_n) = \int_{t_n}^{t_{n+1}} f(t,\, y(t))dt$

on déduit par soustraction :

$$e_{n+1} = e_n + h[\Phi(t_n,\, y_n\, ;\, h) - \frac{1}{h}\int_{t_n}^{t_{n+1}} f(t,\, y(t))dt]$$

$$= e_n + h[\Phi(t_n,\, y_n\, ;\, h) - \Phi(t_n,\, y(t_n)\, ;\, h)]$$

$$+ \quad h[\Phi(t,\, y(t_n)\, ;\, h) - \frac{1}{h}\int_{t_n}^{t_{n+1}} f(t,\, y(t))dt] \;.$$

Utilisant (ii) pour le 1er crochet et (iii) pour le 2e , on en déduit que :

(2,10) $\quad |e_{n+1}| \leqslant |e_n| + h\,L_1\,|e_n| + L_2\,h^{p+1}$.

2/ Pour simplifier l'écriture posons

$$|e_n| = \varepsilon_n \;.$$

Alors (2,10) donne

(2,11) $\quad \varepsilon_{n+1} \leqslant a\,\varepsilon_n + b$

avec (2,12) $\;a = (1+h\,L_1),\; b = L^2 h^{p+1}$.

(2,11) est une inégalité récurrente linéaire. On en déduit donc :

(2,13) $\quad \varepsilon_n \leqslant a^n \varepsilon_o + b\,\dfrac{a^n - 1}{a - 1}$.

Mais $\varepsilon_o = 0$ et

$$\frac{a^n - 1}{a - 1} = \frac{1}{hL_1}\,[e^{n\,\log(1+hL_1)} - 1] \leqslant \frac{1}{hL_1}\,[e^{nhL_1} - 1] \;;$$

or $n\,h = t_n - a$, donc :

(2,14) $\quad \varepsilon_n \leqslant \dfrac{L_2}{L_1}\,(e^{(t_n - a)L_1} - 1)$

Corollaire 2.1. Sous les hypothèses (i), (ii), (iii), la méthode (2,4) est convergente et l'erreur de discrétisation est $O(h^p)$.

Remarque 2.1. Supposons que l'on ait un renseignement plus précis que (iii) :

$$(2,16)\qquad \Phi(t,\ y(t)\ ;\ h) - \frac{y(t+h)-y(t)}{h} = h^p\ \varphi(x) + O(h^{p+1})$$

φ étant continue non nulle dans $[a,b]$. (p est alors l'ordre exact ; la conclusion du corollaire 2.1. est en effet que l'ordre est au moins p).

On peut alors démontrer ceci : soit $V(t)$ la solution de l'équation

$$(2,17)\qquad V'(t) = \frac{\partial f}{\partial y}(t,\ y(t))\ V(t) + \varphi(x)$$

avec

$$(2,18)\qquad V(a) = 0\ .$$

Alors

$$(2,19)\qquad e_n = h^p\ V(t_n) + O(h^{p+1})\ .$$

Naturellement V dépend de l'inconnue y. Mais on peut aussi calculer V en utilisant les approximations de y.

2.4) Exemples : retour aux méthodes d'Euler.

1/ Dans le cas (2,5)

$$\Phi(t,\ y(t)\ ;\ h) - \frac{y(t+h) - y(t)}{h} = f(t,\ y(t)) - \frac{y(t+h) - y(t)}{h}$$

$$= y'(t) - \frac{y(t+h) - y(t)}{h} = -\frac{h}{2}\ y''(t) + O(h^2)$$

d'où d'après le corollaire 2.1. la convergence de la méthode d'Euler (avec l'erreur $O(h)$).

2/ Dans le cas (2,6)

$$\Phi(t,\ y(t)\ ;\ h) = f(t + \tfrac{1}{2}h\ ,\ y(t) + \frac{h}{2} f(t,\ y(t)))$$

$$= f(t,\ y(t)) + \frac{h}{2}\frac{\partial f}{\partial t}(t,\ y(t)) + \frac{h}{2} f(t,\ y(t))\frac{\partial f}{\partial y}(t,\ y(t)) + O(h^2)$$

$$= y'(t) + \frac{h}{2}\ y''(t) + O(h^2)\ .$$

D'où $$\Phi(t,\ y(t)\ ;\ h) - \frac{y(t+h) - y(t)}{h} = O(h^2)\ .$$

Il y a donc effectivement amélioration par rapport à 1/.

2.5) Construction de $\Phi(t,\ y\ ;\ h)$ par utilisation de la série de Taylor.

Un moyen systématique de contruire des $\Phi(t,\ y\ ;\ h)$ satisfaisant à (iii).

Proposition 2.1. est évidemment de développer $\dfrac{y(t+h) - y(t)}{h}$ en série de Taylor, en y remplaçant ensuite y' par $f(t,\ y(t))$ etc... .

On a : $$\frac{y(t+h) - y(t)}{h} = y'(t) + \frac{h}{2}\, y''(t) + \dots + \frac{h^{p-1}}{(p-1)!}\ y^{(p)}(t) + O(h^p)\ .$$

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

d'où

$$(2{,}20)\quad y''(t) = \frac{\partial f}{\partial t}(t,\ y(t)) + y'(t)\frac{\partial f}{\partial y}(t,\ y(t)) = \frac{\partial f}{\partial k}(t,\ y(t)) + f(t,y)\frac{\partial f}{\partial y} f(t,y)$$

et ainsi de suite. Si l'on pose :

$$(2{,}21)\qquad y^{(k)}(t) = {}_{k}(t,\ y(t))$$

$$y^{(k+1)}(t) = \frac{d\,\mathcal{F}_k}{dt} = \frac{\partial\,\mathcal{F}_k}{\partial t}(t,\ y(t)) + \frac{\partial\,\mathcal{F}_k}{\partial y}\ (t,\ y(t)).y'(t)$$

$$= \frac{\partial\,\mathcal{F}_k}{\partial t}\ (t,\ y(t)) + \frac{\partial\,\mathcal{F}_k}{\partial y}\ (t,\ y(t)).f(t,\ y(t))$$

$$= \mathcal{F}_{k+1}\ (t,\ y(t))\ .$$

On peut alors prendre

$$(2{,}22)\qquad \Phi(t,\ y\ ;\ h) = f(t,\ y) + \frac{h}{2}\ \mathcal{F}_1(t,\ y) + \dots + \frac{h^{p-1}}{(p-1)!}\ \mathcal{F}_{p-1}(t,\ y)\ .$$

Avec le choix (2,22) la formule (2,4) donne une méthode convergente avec erreur de discrétisation en $O(h^p)$.

2.6) Construction de $\Phi(t,\ y\ ;\ h)$ par utilisation indirecte de la série de Taylor ; méthode de Runge-Katta.

Le point 2/ de 2.4) donne (à propos de la méthode d'Euler améliorée) un usage différent (et moins direct) de la méthode de Taylor ; cette possibilité a été systématiquement étudiée par Runge-Katta. Considérons ici un cas simple mais typique.

Par analogie avec (2,6) on cherche $\Phi(t, y ; h)$ sous la forme :

(2,23) $$\Phi(t, y ; h) = a_1 f(t,y) + a_2 f(t + p_1 h, y + p_2 h f(t,y)),$$

les constantes a_i , p_i étant à choisir "au mieux".

La condition (i), proposition 2.1. donne :

(2,24) $$a_1 + a_2 = 1 .$$

Développons, en remplaçant y par $y(t)$ et utilisant $y'(t) = f(t, y(t))$:

$$\Phi(t,y(t) ; h) = f(t,y(t)) + a_2 h[p_1 \frac{\partial f}{\partial t}(t,y) + p_2 f(t,y) \frac{\partial f}{\partial y}(t,y)] + O(h^2) ,$$

identifiant à

$$\frac{y(t+h) - y(t)}{h} = y'(t) + \frac{h}{2} y''(t) + O(h^2)$$

$$= f(t,y(t)) + \frac{h}{2} [\frac{\partial f}{\partial t}(t,y) + f \frac{\partial f}{\partial y} (t,y)] + O(h^2) .$$

On voit que (2,23) définit une méthode convergente, en $O(h^2)$, si,outre (2,24) on a :

$$a_2 p_1 = \frac{1}{2} \quad , \quad a_2 p_2 = \frac{1}{2} .$$

Si donc l'on pose : $a_2 = \alpha \neq 0$, il vient :

$$a_1 = 1 - \alpha , \quad p_1 = \frac{1}{2\alpha} \quad , \quad p_2 = \frac{1}{2\alpha} \text{ , d'où}$$

(2,25) $$\Phi(t,y ; h) = (1-\alpha)f(t,y) + \alpha f(t + \frac{h}{2\alpha} , y + \frac{h}{2\alpha} f(t,y)) .$$

Pour ce choix de Φ , la méthode (2,24) est dite de Runge-Katta (simplifiée).

<u>Exemple</u> 2.1. Si l'on prend $\alpha = 1$, on retrouve (2,6).

<u>Remarque</u> 2.2. Il est maintenant facile de poser systématiquement le problème.

On introduit

$$k_o = f(t,y)$$

$$k_1 = f(t + \lambda_1 h,\ y + \mu_{10}\ h\ k_o)$$

$$k_2 = f(t + \lambda_2 h,\ y + \mu_{20}\ h\ k_o + \mu_{21}\ h\ k_1)$$

$$\dots\dots \quad \dots\dots \quad \dots\dots$$

$$k_p = f(t + \lambda_p h,\ y + \mu_{p_o}\ h\ k_o + \dots + \mu_{p,p-1}\ h\ k_{p-1})$$

et on cherche à déterminer les $a_i,\ \lambda_i,\ \mu_i$ de façon que

$$(2{,}26) \qquad \Phi(t,\ y\ ;\ h) = a_o k_o + a_1 k_1 + \dots + a_p k_p$$

de façon que l'ordre de l'erreur soit le plus élevé possible.

Exemple 2.2. On trouvera par exemple

$$(2{,}27) \qquad \Phi(t,\ y\ ;\ h) = \frac{1}{6}\left[k_o + 2k_1 + 2k_2 + k_3\right]$$

avec

$$k_1 = f(t + \tfrac{1}{2}\ h,\ y + \tfrac{1}{2}\ h\ k_o)$$

$$k_2 = f(t + \tfrac{1}{2}\ h,\ y + \tfrac{1}{2}\ h\ k_1)$$

$$k_3 = f(t + h\ ,\ y + h\ k_2)\ .$$

On trouve que cette méthode est d'ordre 4 .

Remarque 2.3. On peut utiliser des processus d'accélération de convergence ; cf. chapitre I § C (3).

Remarque 2.4. Dans les équations différentielles intervenant dans les applications, on sait souvent que la solution est régulière dans un intervalle (approximativement connu) puis irrégulière dans un autre intervalle. Il est alors peu naturel d'utiliser le même pas h tout au long du calcul. On introduit alors des pas variables, t_{n+1} étant défini à partir de t_n par :

$$t_{n+1} = t_n + \theta(t_n)\, h$$

la fonction $t \mapsto \theta(t)$ étant continue et telle que $\theta(t) \geqslant \theta_o > 0$. On remplace alors (2,4) par

$$y_{n+1} = y_n + \theta(t_n)\, h\, \Phi(t_n, y_n\, ;\, \theta(t_n)\, h)\, .$$

3 - Méthodes à plusieurs pas. Utilisation des formules de quadrature.

3.1) Formule de Newton.

1/ On a rappelé au chapitre I § A, la notion de polynôme d'interpolation, les x_j étant des points distincts croissants sur $[a,b]$; $0 \leqslant j \leqslant q$, c'est le polynôme P tel que :

$$(3,1) \quad \begin{cases} P(x_j) = f(x_j) = f_j & (f \text{ donnée}) \text{ pour } 0 \leqslant j \leqslant q \\ d^o\, P \leqslant q \end{cases}$$

Lorsque les x_j sont uniformément répartis sur $[a,b]$

$$(3,2) \qquad x_j = a + j\, h \quad , \quad h = \frac{b-a}{N} \, .$$

On a vu qu'on pouvait donner une expression simple de P en fonctions des différences progressives de la suite $\{f(x_j)\}$ (Cf. Chapitre I, § A proposition 3.4.).

On peut donner une expression analogue de P en utilisant cette fois les différences régressives de la suite $\{f(x_j)\}$.

2/ Opérateur $\bar{\nabla}$.

Soit $\{f_j\}$ une suite donnée. On pose :

$$(3,3) \qquad \bar{\nabla} f_j = f_j - f_{j-1} \quad ;$$

$\bar{\nabla}$ est l'opérateur de différence régressive. On définit ses puissances successives $\bar{\nabla}^2, \bar{\nabla}^3, \ldots$; par définition $\bar{\nabla}^0$ = identité : $\bar{\nabla}^0 f_j = f_j$.

On trouve (par récurrence comme au chapitre I) :

$$\bar{\nabla}^q f_j = \sum_{m=0}^{q} (-1)^m \binom{q}{m} f_{j-m} \quad .$$

On vérifiera que :

$$f_{j-q} = \sum_{m=0}^{q} (-1)^m \binom{q}{m} {}^{-m} f_j \ .$$

Remarque 3.1. Si f_j n'est pas défini pour tout j , mais seulement pour $j \geqslant 0$ (où $0 \leqslant j \leqslant N$), alors $\bar{\nabla} f_j$ n'est défini que pour $j \geqslant 1$, $\bar{\nabla}^2 f_j$ pour $j \geqslant 2$ etc... .

3/ Formule de Newton.

Une notation d'abord ; q et m étant entier, on sait que :

$$\binom{q}{m} = \frac{q(q-1)\ldots(q-m+1)}{1.2\ldots m} \quad .$$

Cette définition s'étend au cas où $q \in \mathbb{R}$ (ou $\in \mathbb{C}$) quelconque. Ceci posé :

Proposition 3.1. Le polynôme $P(x)$ qui vérifie :

$$(3,6) \qquad \begin{cases} d^o \, P \leqslant q \\ P(x_j) = z_j \quad , \quad p-q \leqslant j \leqslant p \qquad (x_j = a + jh) \end{cases}$$

s'exprime par la formule :

$$P(x) = \sum_{m=0}^{q} (-1)^m \binom{-s}{m} {}^{-m} z_p$$

où : (3,8) $\qquad s = \dfrac{x-x_p}{h}$.

Démonstration.

Se reportant à la définition de $\binom{q}{m}$ pour $q \in \mathbb{R}$ et avec (3,8), on voit que $\binom{-s}{m}$ est un polynôme en x de degré m et donc que le 2^e membre de (3,7) est un polynôme de degré $\leqslant q$.

Puis pour $0 \leqslant r \leqslant q$:

$$P(x_{p-r}) = \sum_{m=0}^{q} (-1)^m \binom{r}{m} \bar{\nabla}^m z_p \ .$$

Mais $\binom{r}{m} = 0$ si $r < m$; donc

$$P(x_{p-r}) = \sum_{m=0}^{r} (-1)^m \binom{r}{m} \bar{\nabla}^m z_p \ ,$$

et d'après (3,5) ceci vaut z_{p-r}.

3.2) <u>Utilisation de la formule de quadrature</u>

Soit $y(t)$ solution de (1,1) :

(3,9) $\qquad y'(t) = f(t, y(t)).$

On opère comme pour (1,6) en généralisant un peu. On déduit de (3,9) pour

$$t,\ t+k \in [a,b]$$

(3,10) $\qquad y(t+k) - y(t) = \displaystyle\int_t^{t+k} f(\xi, y(\xi))d\xi \ .$

On calcule approximativement le deuxième membre de (3,10) en remplaçant $f(\xi, y(\xi))$ par <u>le</u> polynôme $P(f\ ;\ \xi) = P(\xi)$ tel que :

(3,11) $\qquad \begin{cases} d^o\ P \leqslant q \\ P(t_j) = f(t_j, y_j) \quad p-q \leqslant j \leqslant p \end{cases}$

y_j ayant déjà été calculé ou que l'on est eventuellement en train de calculer si $j = p$. On pose :

(3,12) $\qquad f_j = f(t_j, y_j)$.

Alors d'après la proposition 3.1.

(3,13) $\qquad P(x) = \displaystyle\sum_{m=0}^{q} (-1)^m \binom{-s}{m} \bar{\nabla}^m f_p \ , \quad s = \frac{x-x_p}{h}$.

On a alors le choix de t, $t+k$ et q.

<u>Laissant</u> q <u>arbitraire pour l'instant</u>, les choix classiques sont :

$$(3,14)\quad \begin{cases} t = t_p \ , & t+k = t_{p+1} \ ; \text{ méthode d'Adams-Bashforth ;} \\ t = t_{p-1} \ , & t+k = t_p \ ; \text{ méthode d'Adams-Moulton ;} \\ t = t_{p-1} \ , & t+k = t_{p+1} \ ; \text{ méthode de Nyström ;} \\ t = t_{p-2} \ , & t+k = t_p \ ; \text{ méthode de Milne-Simpson ;} \end{cases}$$

⟶ La méthode d'Adams-Bashforth donne :

$$(3,15)\qquad y_{p+1} - y_p = h \sum_{m=o}^{q} \gamma_m \bar{\nabla}^m f_p$$

où

$$(3,16)\qquad \gamma_m = (-1)^m \int_0^1 \binom{-s}{m} ds \ .$$

En effet avec le choix indiqué dans (3,14), la formule (3,10) donne :

$$\gamma_m = \frac{1}{h} \int_{t_p}^{t_{p+1}} (-1)^m \binom{-\frac{\xi - t_p}{h}}{m} d\xi$$

d'où (3,16) en posant $\xi - t_p = hs$.

⟶ La méthode d'Adams-Moulton donne :

$$(3,17)\qquad y_p - y_{p-1} = h \sum_{m=o}^{q} \gamma_m^* \bar{\nabla}^m f_p$$

où (3,18) : $\gamma_m^* = (-1)^m \int_{-1}^{0} \binom{-s}{m} ds$.

⟶ La méthode Nyström donne :

$$(3,19)\qquad y_{p+1} - y_{p-1} = h \sum_{m=o}^{q} \chi_m \nabla^m f_p \ .$$

où

$$(3,20)\qquad \chi_m = (-1)^m \int_{-1}^{+1} \binom{-s}{m} ds \ .$$

⟶ La méthode Milne-Simpson donne :

$$(3,21)\qquad y_p - y_{p-2} = h \sum_{m=o}^{q} \chi_m^* \nabla^m f_p$$

où

$$(3,22) \qquad \chi_m^* = (-1)^m \int_{-2}^{0} \binom{-s}{m} \, ds \; .$$

Les problèmes sont maintenant les suivants :

a) un problème technique : trouver des procédés de calcul commode (par récurrence) des coefficients γ_m, γ_m^* ... ; nous allons donner un exemple au point suivant.

b) un problème de comparaison des méthodes, leur intérêt (éventuel), le choix de q ... etc. Ceci fait l'objet du n° 4 ci-dessous.

3.3) <u>Fonctions génératrices</u>.

Pour établir des relations de récurrence (éventuelles) entre les γ_m, on associe à γ_m la fonction

$$(3,23) \qquad \Gamma(t) = \sum_{m=0}^{\infty} \gamma_m \, t^m \qquad t \in \mathbb{C} \; ,$$

(sous réserve de vérification de convergence).

<u>Il s'agit là d'une méthode générale souvent utile</u>.

Remplaçant γ_m par sa valeur (3,16) dans (3,23) il vient :

$$\Gamma(t) = \sum_{m=0}^{\infty} (-t)^m \int_0^1 \binom{-s}{m} \, ds \; ;$$

on peut vérifier qu'il est loisible, pour $|t| < 1$ de commuter la série et l'intégrale, d'où :

$$\Gamma(t) = \int_0^1 \sum_{m=0}^{\infty} (-t)^m \binom{-s}{m} \, ds = \int_0^1 (1-t)^s ds \; ,$$

d'où

$$(3,24) \qquad \Gamma(t) = \frac{-t}{(1-t)\log(1-t)} \qquad |t| < 1 \; .$$

Naturellement, un développement en série de (3,24) reconduirait purement et simplement à (3,23) ; mais écrivons (3,24) sous la forme

$$(3,24 \text{ bis}) \qquad -\frac{1}{t}\log(1-t)\,\Gamma(t) = \frac{1}{1-t}$$

et utilisons les développements :

$$\frac{1}{1-t} = \sum_{n\geq 0} t^n \,,$$

$$-\frac{1}{t}\log(1-t) = \sum_{n\geq 0} \frac{t^n}{n+1}$$

portant dans (3,24 bis) on obtient l'identité :

$$(3,25) \qquad \Big(\sum_{n\geq 0} \frac{t^n}{n+1}\Big)\Big(\sum_{m\geq 0} \gamma_m t^m\Big) = \sum_{n\geq 0} t^n \,,$$

d'où les relations (par identification) :

$$(3,26) \qquad \gamma_m + \frac{1}{2}\gamma_{m-1} + \frac{1}{3}\gamma_{m-2} + \dots + \frac{1}{m+1}\gamma_0 = 1 \,, \quad m = 0, 1, 2, \dots$$

d'où la possibilité de calculer les γ_m par récurrence.

<u>Remarque</u> 3.2. Soient Γ^*, χ, χ^* les fonctions génératrices de γ_m^*, χ_m, χ_m^* définies par (3,18), (3,20), (3,22) ; on trouve :

$$\Gamma^*(t) = -t\,\frac{1}{\log(1-t)} \,, \quad |t| < 1$$

$$\chi(t) = -t\,\frac{2-t}{1-t}\,\frac{1}{\log(1-t)} \,, \quad |t| < 1$$

$$\chi^*(t) = -t\,(2-t)\,\frac{1}{\log(1-t)} \,, \quad |t| < 1 \,.$$

4 - Etude générale des méthodes à plusieurs pas. Stabilité.

4.1) Formulation générale.

Tous les schémas introduits au n° 3 entrent dans le type suivant :

$$(4,1) \quad \alpha_k y_{n+k} + \alpha_{k-1} y_{n+k-1} + \dots + \alpha_0 y_n = h[\beta_k f_{n+k} + \beta_{k-1} f_{n+k-1} + \dots + \beta_0 f_n]$$

où les α_i, β_i sont des nombres réels donnés,

$$f_j = f(t_j, y_j) \qquad t_j = a + jh \,, \quad h = \frac{b-a}{N} \,.$$

On supposera :

(4,2) $\alpha_k \neq 0$

et (4,3) $|\alpha_o| + |\beta_o| \neq 0$.

(De sorte que y_n intervient vraiment dans (4,1)).

<u>Le schéma (4,1) est le schéma général à k pas.</u>

<u>Remarque</u> 4.1. Comme on a déjà signalé (remarque 1.1.) il y a une <u>difficulté au démarrage</u> puisque l'on ne peut espérer obtenir y_{n+k} par (4,1) sans connaître $y_{n+k-1}, \ldots, y_n$. Nous reviendrons sur ce point.

. Méthode explicite. Si $\beta_k = C$, (4,1) est explicite cf. déjà (1,2) n° 1 :

. Méthode implicite. Si $\beta_k \neq 0$ (4,1) est implicite ; on obtient (éventuellement) y_{n+k} par la résolution de :

(4,4) $\alpha_k \, y_{n+k} - h \, \beta_k \, f(t_{n+k}, y_{n-k}) = \text{donné} = e$.

Posons un instant $y_{n+k} = y$. $\frac{h \, \beta k}{\alpha_k} \, f(t_{n+k}, y) = F(y)$.

Alors (4,4) s'écrit :

(4,5) $y = F(y) + d = G(y)$, $d = \frac{c}{\alpha_k}$.

Utilisant le fait que f est L-lipschitzienne, on a :

$$|G(y) - G(y^*)| \leqslant h \left|\frac{\beta_k}{\alpha_k}\right| L \, |y-y^*|$$

d'où (cf. cours d'Analyse) :

(4,6) $\left\{ \begin{array}{l} \text{si } \beta_k \neq 0 \text{ , la formule (4,1) définit } y_{n+k} \text{ de} \\ \text{façon unique si } h \left|\frac{\beta_k}{\alpha_k}\right| L > 1 \text{ .} \end{array} \right.$

Il faudra donc choisir h assez petit pour que cette condition ait lieu.

Les problèmes sont maintenant les suivants : comment choisir "au mieux" les coefficients α_j , β_j dans (4,1) ?

4.2) <u>Définition de la convergence</u>.

Introduisons les polynômes (4,7) $\begin{cases} \alpha(\lambda) = \alpha_o + \alpha_1\lambda + \ldots + \alpha_k\lambda^k \\ \beta(\lambda) = \beta_o + \beta_1\lambda + \ldots + \beta_k\lambda^k \end{cases}$

et <u>désignons par</u> (α, β) <u>le schéma (4,1)</u>.

A toute fonction $f(x,y) = f$ vérifiant (1,3) on associe :

a) le problème de Cauchy $y'(t) = f(t, y(t))$, $y(t_o) = y_o$.

b) le problème "approché" (4,1) [avec (4,6) si $\beta_k \neq 0$] avec les données initiales.

(4,8) $\qquad y_j = \eta_j(h) \qquad 0 \leqslant j \leqslant k-1$

les η_j satisfaisant à la condition (minimum) :

(4,9) $\qquad \eta_j(h) \to y_o$ lorsque $h \to 0$, $\quad 0 \leqslant j \leqslant k-1$.

N.B. On ne prend pas nécessairement y_o coïncidant avec la donnée de Cauchy.

<u>Définition</u> 4.1. Le schéma (α,β) est dit <u>convergent</u> si, sous la conditions précédentes (et pour <u>toute</u> f), on a :

(4,10) $\qquad y_n - y(t_n) \to 0$ lorsque $h \to 0$, $t_n = a + nh$ fixé $\leqslant b$.

Z N.B. La convergence, dans (4,10) n'est pas uniforme en f .

Notre but est maintenant de donner des conditions nécessaires et suffisantes sur les polynômes α, β pour qu'il y ait convergence.

On va commencer par l'étude (importante en elle-même) des équations aux différences.

4.3.) <u>Equations linéaires aux différences. Leur stabilité</u>.

Sur l'espace S des suites $y = \{y_n\}$, $n \geqslant 0$, on considère <u>l'opérateur</u>

linéaire Λ_α défini par :

$$(\Lambda_\alpha y)_n = \alpha_k y_{n+k} + \alpha_{k-1} y_{n+k} + \dots + \alpha_o y_n \quad , \qquad a_k \neq 0$$

$$\Lambda_\alpha y = \{(\Lambda_\alpha y)_n\} \ .$$

Définition 4.2. L'opérateur Λ_α -et, par abus de langage, le polynôme α - est dit stable, si toute suite $\{y_n\}$ solution de :

(4,11) $\quad \Lambda_\alpha y = 0$ est bornée.

Montrons là

Proposition 4.1. La condition nécessaire et suffisante de stabilité de l'opérateur Λ_α est que les deux conditions suivantes aient lieu :

(i) les racines du polynôme α sont dans le disque $|\lambda| < 1$

(ii) les racines éventuellement situées sur ce cercle $|\lambda| = 1$ sont simples.

Démonstration.

1/ Les solutions de (4,11) forment un sous-espace de S de dimension k .

Soit en effet $E^j = \{E_n^j\}$, $0 \leqslant j \leqslant k-1$ la solution de :

$$(4,12) \qquad \begin{cases} \Lambda_\alpha E^j = 0 \\ E_m^j = \delta_m^j \quad , \quad 0 \leqslant m \leqslant k-1 \end{cases}$$

Alors y solution quelconque de (4,11) s'exprime par

$$y = \sum_{j=0}^{k-1} y_j E^j \quad \text{d'où le résultat.}$$

2/ Solutions particulières.

La suite $\{n^r \lambda^n\}$, r entier, $\lambda \in \mathbb{C}$ est solution de (4,11) si et seulement si λ est racine d'ordre $r+1$ de α (faire le calcul de $(\Lambda_\alpha y_n)$,

$y_n = n^r \lambda^n$) et alors $\{\lambda^n\}$, $\{n\, \lambda^n\}, \ldots, \{n^r\, \lambda^n\}$ sont solutions linéairement indépendantes.

Soient alors λ_j les racines de $\alpha(\lambda) = 0$, λ_j étant d'ordre m_j, $0 \leqslant j \leqslant \nu$, $\sum_{j=1}^{\nu} m_j = k$. Les suites

$$(4,13) \qquad \{n^{r_j}\, \lambda_j^n\}, \qquad 0 \leqslant r_j \leqslant m_j, \quad 1 \leqslant j \leqslant \nu$$

sont linéairement indépendantes, en nombre k, et donc forment une base de l'espace des solutions de (4,11).

3/ La condition nécessaire et suffisante de stabilité est alors que <u>toute</u> suite (4,13) soit bornée, d'où la proposition.

<u>Remarque</u> 4.3. Ecrivons explicitement <u>la</u> solution de l'équation non homogène :

$$(4,14) \qquad \alpha_k\, y_{n+k} + \ldots + \alpha_o\, y_n = d_{n+k}, \qquad n \geqslant 0$$

$$(4,15) \qquad y_j = \eta_j, \qquad 0 \leqslant j \leqslant k.$$

Introduisons la solution élémentaire $E = \{E_n\}$, $n \in \mathbb{Z}$, avec :

$$E_n = E_n^{k-1} \quad \text{si} \quad n \geqslant 0$$

$$E_n = 0 \quad \text{si} \quad n < 0.$$

Alors
$$\sum_{j=o}^{k} \alpha_j\, E_{n+j} = 0 \quad \text{si} \quad n \geqslant 0$$
$$= \alpha_k \quad \text{si} \quad n = -1$$
$$= 0 \quad \text{si} \quad n < -1 \quad [n + j \leqslant k].$$

Donc :

$$\sum_{j=o}^{k} \alpha_j\, E_{n+j} = \begin{cases} 0 & \text{si } n \neq -1 \\ \alpha_k & \text{si } n = -1 \end{cases}$$

Définissons alors $E \otimes d$ par :

$$(4,17) \qquad (E\otimes d)_n = \sum_{m=0}^{\infty} E_{n-m-1}\, d_{m+k} \quad \text{sin} > 0 .$$

Alors, comme $E_{n-m-1} = 0$ dès que $n-m-1 < k-1$ i.e. $m > n-k$

$$(E\otimes d)_n = \sum_{m=0}^{n-k} E_{n-m-1}\, d_{m+k} \quad \text{si} \quad n \geqslant k$$

$$= 0 \qquad\qquad \text{si} \quad n < k .$$

La solution de (4,14), (4,15) est alors fournie par :

$$(4,18) \qquad y = \sum_{j=0}^{k-1} \eta_j\, E^j + \frac{1}{\alpha_k}\, E\otimes d .$$

En effet $(E\otimes d)_n = 0$ si $0 \leqslant n \leqslant k-1$ est la seule chose à vérifier est

que pour $h \geqslant 0$

$$\xi_n = \sum_{j=0}^{k} \alpha_j\, (E\otimes d)_{n+j} \quad \text{est égal à} \quad \alpha_k\, d_{n+k}$$

or
$$\xi_n = \sum_{m=0}^{\infty} d_{m+k} \left(\sum_{j=0}^{k} \alpha_j\, E_{n+j-m-1}\right),$$

et d'après (4,16)

$$\sum_{j=0}^{k} \alpha_j\, E_{n+j-m-1} = 0 ,$$

pour tout $n \neq m$ et vaut α_k si $n = m$ d'où (4,18).

<u>Proposition</u> 4.2. Si le polynôme α est stable, il existe une constante K_α

telle que, y_n étant la solution de (4,14), (4,15), on a :

$$(4,19) \qquad |y_n| \leqslant K_\alpha \left(\sum_{j=0}^{k-1} |\eta_j| + \sum_{m=0}^{n-k} |d_{m+k}|\right), \qquad (n \geqslant k) .$$

<u>Démonstration</u> On peut en effet expliciter (4,18) sous la forme

$$y_n = \sum_{j=0}^{k-1} \eta_j\, E_n^j + \frac{1}{\alpha_k} \sum_{m=0}^{n-k} E_{n-m-1}^{k-1}\, d_{m+k}$$

d'où (4,19) grâce au fait que $\{E^j\}$, $0 \leqslant j \leqslant k-1$, est bornée.

4.4) <u>Conditions nécessaires de convergence</u>.

<u>Proposition</u> 4.3. Pour que le schéma (α, β) soit convergent, il faut que le

polynôme α soit stable.

<u>Démonstration.</u>

Si le polynôme α n'est pas stable il existe une suite $Z = \{Z_n\}_n$ vérifiant

* $\Lambda_\alpha Z = 0$

* $\lim\limits_{n\to\infty} Z_n = \infty$.

Considérons un point x et des valeurs de h telles que :

$$h_\lambda = \frac{x}{\lambda} \qquad \lambda \in \mathbb{N} .$$

Considérons la suite $y_n = \dfrac{Z_n}{\sqrt{|Z_\lambda|}}$ lorsque h tend vers zéro

$$y_i \to 0 \qquad \forall : i = 0, 1, \dots, k-1 .$$

Mais $y_\lambda = \sqrt{|Z_\lambda|}$ tend vers l'infini.

La valeur de la "solution approchée" au point x ne tend donc pas vers zéro.

Si on applique ceci au problème de Cauchy

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

on trouve une contradiction. La méthode n'étant pas convergente.

<u>Démonstration</u>. (suite)

1/ Considérons l'équation particulière

$$y' = 0 \ , \ y(0) = 1$$

de solution $y(x) = 1$. Le schéma (α, β) se réduit à :

$$(4,22) \qquad \sum_{j=0}^{k} \alpha_j \, y_{n+j} = 0 \quad ; \quad y_j = \eta_j(h) \qquad 0 \leqslant j \leqslant k-1$$

avec $\eta_j(h) \to 1$.

Prenons en particulier

$$\eta_j(h) = 1 \ , \qquad 0 \leqslant j \leqslant k-1 \ .$$

La solution y_n ne dépend pas de h et la condition de convergence se traduit par $y_n \to 1$ lorsque $n \to \infty$; passant à la limite en n dans (4,22), il vient :

$$\sum_{j=0}^{k} \alpha_j = \alpha(1) = 0 \ .$$

2/ D'après la proposition 4.3., le polynôme α doit être stable. Donc la racine 1 est nécessairement simple (proposition 4.1.), donc :

$$\alpha'(1) \neq 0 \ .$$

3/ Considérons maintenant l'équation particulière :

$$y' = 1 \quad , \quad y(0) = 0$$

de solution $y(t) = t$. Le schéma (α, β) se réduit à :

$$(4,23) \qquad \sum_{j=0}^{k} \alpha_j \, y_{n+j} = h \, (\beta_k + \beta_{k-1} + \ldots + \beta_o) = h \, \beta(1) \ .$$

Choisissons

$$(4,22) \qquad \eta_j(h) = \frac{\beta(1)}{\alpha'(1)} \, j \, h \qquad \text{(loisible d'après 2/).}$$

Alors la solution de (4,23), (4,24) est :

$$y_n = n\, h \frac{\beta(1)}{\alpha'(1)}$$

et $$y(t_n) = n\, h \ ;$$

la convergence de y_n vers $y(t_n)$ équivaut donc à $\frac{\beta(1)}{\alpha'(1)} = 1$, d'où la proposition.

On est maintenant en mesure de démontrer le premier résultat essentiel de la théorie des schémas (α, β). C'est l'objet du 4.5).

4.5) Condition nécessaire et suffisante de convergence.

Théorème 4.1. La condition nécessaire et suffisante pour que le schéma (α, β) soit convergent est que les deux conditions suivantes aient lieu :

(i) le polynôme α est stable (cf. proposition 4.1.)

(ii) $\alpha(1) = 0$, $\alpha'(1) = \beta(1)$ $(\neq 0)$.

Démonstration.

1/ La nécessité de (i) et (ii) résulte des propositions 4.3. et 4.4.

2/ Suffisance des conditions.

Formons une inéquation aux différences satisfaites par $e_n = y_n - y(t_n)$ [comparer à la démonstration de la proposition 2.1.]. Posons :

$$\alpha_k\, y(t_{n+k}) + \ldots + \alpha_o\, y(t_n) - h[\beta_k\, f(t_{n+k}, y(t_{n+k}))] + \ldots$$
$$\ldots + \beta_o\, f(t_n, y(t_n)) = \zeta_n \ .$$

Alors

$$(4,25) \qquad \alpha_k\, e_{n+k} + \ldots + \alpha_o\, e_n - h \sum_{j=o}^{k} \beta_j [f_{n+j} - f(t_{n+j}, y(t_{n+j}))] = -\zeta_n$$

Mais $$f_{n+j} - f(t_{n+j}, y(t_{n+j})) = f(t_{n+j}, y_{n+j}) - f(t_{n+j}, y(t_{n+j}))$$
$$= g_{n+j}\, e_{n+j} ,$$

avec

$$(4,26) \qquad |g_{n+j}| \leqslant C_1 \quad \text{(constante)}.$$

Par ailleurs

$$y(t_{n+j}) = y(t_n + j\,h) = y(t_n) + j\,h\,y'(t_n) + o(j\,h)$$
$$\left[\frac{o(h)}{h} \to 0 \quad \text{si} \quad h \to 0\right],$$
$$y'(t_{n+j}) = y'(t_n + j\,h) = y'(t_n) + O(j\,h)$$

d'où en tenant compte de (ii)

$$(4,27) \qquad \zeta_n = o(h)$$

On écrit alors (4,25) sous la forme :

$$(4,28) \quad \begin{cases} \alpha_k\, e_{n+k} + \dots + \alpha_o\, e_n = d_{n+k} \;, & n \geqslant 0 \\ d_{n+k} = h \displaystyle\sum_{j=o}^{k} \beta_j\, g_{n+j}\, e_{n+j} - \zeta_n \;, & \\ e_j = \eta(h) - y(t_j) \;, & 0 \leqslant j \leqslant k-1 \;. \end{cases}$$

Comme $\eta_j(h) \to y(t_o)$ lorsque $h \to 0$, on peut poser :

$$(4,29) \qquad \sum_{o\leqslant j\leqslant k-1} |e_j| = \delta(h), \quad \delta(h) \to 0 \;, \text{ si } \; h \to 0 \;.$$

Comme le polynôme α est stable, la proposition 4.2. donne :

$$|e_n| \leqslant K_\alpha \left(\delta(h) + \sum_{m=o}^{n-k} |d_{m+k}|\right).$$

Or $$|d_{m+k}| \leqslant C_2\, h \sum_{o\leqslant j\leqslant k} |e_{m+j}| + o(h)$$

d'où $$\sum_{m=o}^{n-k} |d_{m+k}| \leqslant C_3\, h \sum_{p=o}^{n} |e_p| + o(1)$$

$[o(1) \to 0$ si $h \to 0]$. Donc :

$$(4,30) \qquad |e_n| \leqslant K_\alpha \, (\delta(h) + o(1) + C_3 \, h \sum_{p=o} |e_p|) \, .$$

Choisissons h assez petit pour que

$1 - C_3 \, h \geqslant \frac{1}{2}$. Alors (4,30) donne

$$|e_n| \leqslant 2 \, K_\alpha \, (\delta(h) + o(1) + C_3 \, h \sum_{p=o}^{n-1} |e_p|)$$

et finalement posant :

$$(4,31) \qquad |e_n| = \varepsilon_n \, .$$

On a :

$$(4,32) \quad \begin{cases} \varepsilon_n \leqslant a(h) + b \, h \displaystyle\sum_{p=o}^{n-1} \varepsilon_p \\ a(h) \to 0 \quad \text{si} \quad h \to 0 \quad , \quad b = \text{constante}. \end{cases}$$

Mais on déduit <u>par récurrence</u> à partir de (4,32) que :

$$(4,33) \qquad \varepsilon_n \leqslant (a(h) + b \, h \, \varepsilon_o) \, (1 + b \, h)^{n-1} \, .$$

Mais $(1 + b\,h)^n = \exp(n \log (1 + b\,h))$ est borné (car $n\,h$ fixé) d'où $\varepsilon_n \to 0$, ce qui démontre le théorème.

<u>Remarque</u> 4.4. <u>Stabilité du schéma</u>.

L'inégalité (4,19) de la proposition 4.2. montre que si le polynôme α est stable, alors de petites variations des données initiales du second membre entrainent de petites variations sur la solution de (4,14), (4,15).

Le schéma (α, β) est dit stable si de petites variations des données initiales (4,8) entrainent de petites variations pour la solution du schéma (4,1) [avec (4,8)]. Plus précisément si :

$$|\eta_j(h)| \leqslant M \; , \quad \exists \, M_1 \quad \text{t.q.} \quad |y_n| \leqslant M_1 \quad (h \to 0 \; , \; n\,h \; \text{fixé} \; M_1 \; \text{constante}).$$

Vérifions brièvement que la stabilité du schéma (α, β) équivaut à la stabilité du polynôme α .

Ecrivons (4,1) de la façon suivante ; notons que [analogue au calcul fait à propos du théorème 4.1. en utilisant 0 au lieu de $y(t_{n+j})$]

$$f_{n+j} = f(t_{n+j}, y_{n+j}) = f(t_{n+j}, y_{n+j}) - f(t_{n+j}, 0) + f(t_{n+j}, 0)$$

entraîne

$$|f_{n+j}| \leqslant L\,|y_{n+j}| + C \quad , \qquad C = \sup|f(t, 0)|$$

de sorte que (4,1) donne :

$$\sum_{j=o}^{k} \alpha_j\, y_{n+j} = g_{n+k} \;,$$

$$g_{n+k}| \leqslant C_1\, h \sum_{j=o}^{k} |y_{n+j}| + C_2\, h \;.$$

On déduit de là, comme au théorème 4.1. et avec la proposition 4.2. l'assertion ci-dessus.

<u>Remarque</u> 4.5. <u>Consistance du schéma</u> (α, β) .

Le schéma (4,1) (où (α, β) est dit <u>consistant</u> si lorsqu'on remplace dans (4,1) y_n par $y(t_n)$ la différence des deux membres est $O(h)$; plus précisément :

$$(4,34)\quad \left\{ \begin{array}{l} \alpha_k\, y(t_{n+k}) + \alpha_{k-1}\, y(t_{n+k-1}) + \ldots + \alpha_o\, y(t_n) \\ - h[\beta_k\, f(t_{n+k}, y(t_{n+k}))] + \ldots + \beta_o\, f(t_n, y(t_n)) = o(h). \end{array} \right.$$

On vérifie sans peine (exercice) que <u>la consistance du schéma</u> (α, β) <u>équivaut à la condition (ii) du théorème</u> 4.1. .

<u>Remarque</u> 4.6. <u>Autre énoncé</u> (équivalent) <u>du théorème</u> 4.1. .

Avec la terminologie introduite dans les remarques 4.4. et 4.5. on peut énoncer le théorème 4.1. sous la forme équivalente suivante :

Théorème 4.1. bis. Le schéma (α, β) est convergent si et seulement si il est stable et consistant.

Remarque 4.7. Tout ce qui a été dit s'étend au cas des systèmes. Il suffit de remplacer les modules par des normes. Tout ceci vaut encore pour des équations différentielles dans des espaces de Banach.

Remarque 4.8. Dans la définition 4.1. de la convergence, nous avons utilisé la convergence ponctuelle. Or (4,33) entraîne la convergence uniforme.

Pour $\varepsilon > 0$, on aura $|y_n - y(t_n)| = \varepsilon$, pour h assez petit et tout n tel que $a + nh \leqslant b$.

4.6) Exemples.

Voyons quels sont les polynômes α et β dans les exemples introduits au n° 3.

Méthode d'Adams-Bashforth

Voici comment trouver aisément les polynômes α et β liés à (3,15) (et même chose pour (3,17) etc...). Si à la suite $\{f_n\}_{n\in\mathbb{Z}}$, on associe la fonction $f = \sum_{n\in\mathbb{Z}} f_n \lambda^{-n}$, l'opérateur de translation $\{f_n\} \to \{f_{n+1}\}$ se traduit sur les fonctions correspondantes par la multiplication par λ.

L'opérateur $\bar{\nabla}$ correspond alors à la multiplication par $(1 - \frac{1}{\lambda})$, et $\bar{\nabla}^m$ à $(1 - \frac{1}{\lambda})^m$.

Dans (3,15) nous posons (pour nous ramener aux notations du n° 4 4.1.) : $p+1 = n+k$; alors $p-q = n$ d'où $q = k-1$ et l'on trouve

$$(4,35) \quad \begin{cases} \alpha(\lambda) = \lambda^k - \lambda^{k-1} \\ \beta(\lambda) = \lambda^{k-1} \left(\sum_{m=0}^{k-1} \gamma_m \left(1 - \frac{1}{\lambda}\right)^m\right) \end{cases}.$$

$\boxed{\text{Méthode d'Adams-Moulton}}$ [cf. (3,11), (3,18)].

$$(4,36) \quad \begin{cases} \alpha(\lambda) = \lambda^k - \lambda^{k-1} \\ \beta(\lambda) = \lambda^k \sum_{m=0}^{k} \gamma_m^* \left(1 - \frac{1}{\lambda}\right)^m \end{cases}.$$

$\boxed{\text{Méthode de Nyström}}$ [cf. (3,19), (3,20)].

$$(4,37) \quad \begin{cases} \alpha(\lambda) = \lambda^k - \lambda^{k-2} \\ \beta(\lambda) = \lambda^{k-1} \sum_{m=0}^{k-1} \chi_m \left(1 - \frac{1}{\lambda}\right)^m \end{cases}.$$

$\boxed{\text{Méthode de Milne Simpson}}$ [cf. (3,21), (3,22)].

$$(4,38) \quad \begin{cases} \alpha(\lambda) = \lambda^k - \lambda^{k-2} \\ \beta(\lambda) = \lambda^k \sum_{m=0}^{k} \chi_m^* \left(1 - \frac{1}{\lambda}\right)^m \end{cases}.$$

Grâce au théorème 4.1. on vérifie sans peine la convergence de ces méthodes.

Remarque 4.9. La méthode (α, β) est explicite (resp. implicite) si $d^o\beta < d^o\alpha$ (resp. $d^o\beta = d^o\alpha$).

Remarque 4.10. La question est maintenant évidemment : comment choisir entre ces diverses méthodes (α, β) ?

Des considérations "techniques" permettent de montrer que les méthodes d'Adams-Bashforth et Adams-Moulton ne sont pas optimales. Les méthodes de Nyström et Milne-Simpson ne sont éventuellement optimales que si $k = 2$. Alors $\alpha(\lambda) = \lambda^2 - 1$ et dans ce cas le polynôme β optimal est $\beta(\lambda)$ $\beta(\lambda) = \frac{1}{3}(\lambda^2 + 4\lambda + 1)$.

Cette méthode, appelée méthode de Milne est convergente d'ordre 4.