

*Andrzej Lasota*

## Sur le choix optimal des points de division dans la méthode de Euler-Cauchy de l'intégration approximative des équations différentielles

1. Nous considérons l'équation différentielle

$$(1) \quad x' = f(t, x).$$

Supposons que la fonction  $f$  soit définie, continue dans l'ensemble

$$\pi: \quad 0 \leq t \leq a, \quad -\infty < x < +\infty$$

et telle qu'il existe dans l'intervalle  $\langle 0, a \rangle$  une et une seule solution  $x_f(t)$  de (1) satisfaisant à la condition initiale

$$(2) \quad x(0) = 0.$$

Soit  $\tau = (t_1, \dots, t_n)$  une suite de points tels que

$$(3) \quad t_0 = 0 \leq t_1 \leq \dots \leq t_n \leq a = t_{n+1}.$$

Dans la méthode de Euler-Cauchy la solution approximative déterminée par la suite  $\tau$  est donnée par les formules suivantes

$$(4) \quad x_{f\tau}(t) = x_i + (t - t_i)f(t_i, x_i), \quad t_i \leq t \leq t_{i+1}$$

où

$$x_0 = 0, \quad x_i = x_{f\tau}(t_i), \quad i = 0, \dots, n.$$

L'erreur maximale

$$m(f, \tau) = \max_{0 \leq t \leq a} |x_f(t) - x_{f\tau}(t)|$$

dépend évidemment de la fonction  $f$  et du choix des points  $t_1, \dots, t_n$ . Pour le second membre de l'équation (1) fixé la fonction  $m(f, \tau)$  est continue dans l'ensemble (3) par rapport à  $\tau$  et admet sa plus petite valeur pour un  $\tau = \tau_0$ . Le problème de déterminer  $\tau_0$  est très difficile même dans le cas où l'équation (1) s'intègre effectivement. Cependant ce problème est d'autant plus important que dans certains cas la suite optimale  $\tau_0$  pour une fonction  $f$  est aussi une suite optimale pour toute une classe de fonctions  $f$ .

2. Lemme. Soit  $x_p(t)$  la solution de l'équation

$$(5) \quad x' = g(t, x), \quad g(t, x) \equiv K + Lt + Mx$$

( $K, L, M$  désignent des constantes non négatives et telles que  $L + KM > 0$ ) satisfaisant à la condition (2) et soit  $x_{gr}(t)$  la solution approximative de cette équation donnée par les formules (4).

La fonction  $x_p(t) - x_{gr}(t)$  est positive et croissante dans l'intervalle  $(0, a)$ . De plus l'expression

$$m(g, \tau) = \max_{0 \leq t \leq a} |x_p(t) - x_{gr}(t)| = x_p(a) - x_{gr}(a)$$

admet sa valeur minimale pour la suite  $\tau_0 = (t_1^0, \dots, t_n^0)$  telle que

$$t_i^0 = \frac{ai}{n+1} \quad i = 1, \dots, n.$$

Démonstration. Posons pour abrégier  $\varphi(t) = x_p(t) - x_{gr}(t)$  et supposons que  $\varphi(t_i) \geq 0$ . Dans l'intervalle  $(t_i, t_{i+1})$  on a

$$\varphi'(t) = g(t, x_p(t)) - g(t_i, x_{gr}(t_i)) = L(t - t_i) + M(x_p(t_i) - x_{gr}(t_i)) + M(x_p(t) - x_p(t_i)).$$

Le premier composant de cette somme est évidemment non négatif (positif si  $L > 0$ ), le second est non négatif par l'hypothèse et le troisième est aussi non négatif (positif si  $KM > 0$ ) parce que la solution  $x_p(t)$  de l'équation (1) est, comme il est facile de vérifier, non-décroissante (croissante si  $KM > 0$ ). Donc, de l'hypothèse  $L + KM > 0$  il vient  $\varphi'(t) > 0$  pour tout  $t \in (t_i, t_{i+1})$  ce qui démontre que de la condition  $\varphi(t_i) \geq 0$  il résulte que la fonction  $\varphi(t)$  est croissante et positive dans l'intervalle  $(t_i, t_{i+1})$ . Comme  $\varphi(0) = x_p(0) - x_{gr}(0) = 0$ , il en vient par récurrence que  $\varphi(t)$  est croissante et positive dans tout l'intervalle  $(0, a)$ .

Comme  $x_p(a)$  ne dépend pas de  $\tau$ , pour achever la démonstration il suffit de montrer que  $x_{gr}(a)$  admet sa valeur maximale pour  $\tau = \tau_0$ . Remarquons d'abord que la fonction  $x_{gr}(t)$  a les propriétés suivantes.

1° Si des suites

$$\tau_1 = (t_1^1, \dots, t_n^1), \quad \tau_2 = (t_1^2, \dots, t_n^2)$$

satisfont aux inégalités (3) et aux conditions

$$t_i^1 = t_i^2, \quad t_{i+1}^1 = t_{i+1}^2, \quad x_{gr}(t_i^1) > x_{gr}(t_i^2),$$

on a

$$x_{gr_2}(t_{i+1}^1) > x_{gr_1}(t_{i+1}^1).$$

En effet, de (4) il vient

$$x_{gr_2}(t_{i+1}^1) - x_{gr_1}(t_{i+1}^1) = (x_{gr_2}(t_i^1) - x_{gr_1}(t_i^1)) (1 + M(t_{i+1}^1 - t_i^1)) > 0.$$

2° Pour  $\tau(u) = (t_1, \dots, t_{i-1}, u, t_{i+1}, \dots, t_n)$ , la fonction  $a(u) = x_{gr(u)}(t_{i+1})$  admet sa valeur maximale stricte pour  $u = (t_{i-1} + t_{i+1})/2$ .

D'après (4) on a

$$\alpha(u) = \beta(u) + (t_{i+1} - u)(K + Lu + M\beta(u))$$

$$\beta(u) = x_{i-1} + (u - t_{i-1})(K + Lt_{i-1} + Mx_{i-1})$$

c'est-à-dire

$$\alpha(u) = -(L + M(K + Lt_{i-1} + Mx_{i-1}))u^2 + \\ + (t_{i+1} + t_{i-1})(L + M(K + Lt_{i-1} + Mx_{i-1}))u + \dots$$

Donc, en vertu de l'hypothèse  $L + KM > 0$ ,  $\alpha(u)$  est un polynôme de degré 2 et il admet sa valeur maximale pour  $u = (t_{i+1} + t_{i-1})/2$ .

En utilisant 1° et 2° il est facile d'achever la démonstration. Supposons que la fonction  $x_{\sigma_i}(a)$  admette sa valeur maximale pour  $\tau_1 = (t_1^1, \dots, t_n^1)$  et admettons qu'il existe un indice  $j$  pour lequel  $t_j^1 \neq a_j/(n+1)$ . Alors, pour un indice  $k$  on a forcément  $t_k^1 \neq (t_{k+1}^1 + t_{k-1}^1)/2$ . En remplaçant  $t_k^1$  par  $t_k^* = (t_{k+1}^1 + t_{k-1}^1)/2$  et en désignant par  $\tau_2$  la suite  $(t_1^1, \dots, t_{k-1}^1, t_k^*, t_{k+2}^1, \dots, t_n^1)$ , de 2° on obtient  $x_{\sigma_2}(t_{k+1}) > x_{\sigma_1}(t_{k+1})$ . De 1° il vient que cette inégalité reste valable pour  $i = k+1, k+2, \dots, n$  d'où l'on tire enfin  $x_{\sigma_2}(a) > x_{\sigma_1}(a)$  ce qui est impossible.

3. Théorème. Soit  $F$  la famille de fonctions définies dans l'ensemble  $\pi$  et satisfaisant aux conditions

$$(6) \quad |f(0, 0)| \leq K, \quad |f(\bar{t}, \bar{x}) - f(t, x)| \leq L|\bar{t} - t| + M|\bar{x} - x|$$

où  $K, L, M$  désignent des constantes non négatives telles que  $L + KM > 0$ .

La fonction

$$\varepsilon(\tau) = \sup_{\substack{0 \leq t \leq a \\ f \in F}} |x_f(t) - x_{f_\tau}(t)|$$

admet dans l'ensemble (3) sa valeur minimale pour  $\tau_0 = (t_1^0, \dots, t_n^0)$  tel que

$$t_i^0 = \frac{ai}{n+1} \quad i = 1, \dots, n.$$

Remarque. Si  $L + KM = 0$  c'est-à-dire si  $L = 0$  et  $KM = 0$ , il est facile de voir que les fonctions  $x_f(t)$  et  $x_{f_\tau}(t)$  sont données par la formule  $x_f(t) = x_{f_\tau}(t) = tf(0, 0)$  et la question envisagée devient banale.

Démonstration. Il est évident que la fonction  $g(t, x) = K + Lt + Mx$  appartient à  $F$ . Par conséquent, en vertu du lemme il suffit de montrer que pour tout  $f \in F$  on a

$$\max_{0 \leq t \leq a} |x_f(t) - x_{f_\tau}(t)| \leq \max_{0 \leq t \leq a} |x_g(t) - x_{g_\tau}(t)|.$$

A cet effet nous allons montrer que pour tout  $f \in F$  on a dans l'intervalle  $\langle 0, a \rangle$  les inégalités suivantes

$$(7) \quad |x_f(t) - x_{f_\tau}(t)| \leq x_g(t) - x_{g_\tau}(t), \quad |x_{f_\tau}(t)| \leq x_{g_\tau}(t).$$

Supposons que ces inégalités soient satisfaites pour  $t = t_i$ .

De (1), (4) et (6) il vient

$$\begin{aligned} |x'_f(t) - x'_{f_r}(t)| &= |f(t, x_f(t)) - f(t_i, x_{f_r}(t_i))| \\ &\leq L(t - t_i) + M|x_f(t) - x_{f_r}(t_i)|, \quad t_i \leq t \leq t_{i+1} \end{aligned}$$

et ensuite

$$\begin{aligned} (8) \quad |x'_f(t) - x'_{f_r}(t)| &\leq L(t - t_i) + M|x_f(t) - x_{f_r}(t)| + M|x_{f_r}(t) - x_{f_r}(t_i)| \\ &= L(t - t_i) + M|x_f(t) - x_{f_r}(t)| + M(t - t_i)|f(t_i, x_{f_r}(t_i))| \\ &\leq L(t - t_i) + M|x_f(t) - x_{f_r}(t)| + M(t - t_i)(K + Lt_i + M|x_{f_r}(t_i)|), \\ & \quad t_i \leq t \leq t_{i+1}. \end{aligned}$$

De la même façon on obtient

$$\begin{aligned} (9) \quad x'_g(t) - x'_{g_r}(t) &= (K + Lt + Mx_g(t)) - (K + Lt_i + Mx_{g_r}(t_i)) \\ &= L(t - t_i) + M|x_g(t) - x_{g_r}(t_i)| + M(x_{g_r}(t) - x_{g_r}(t_i)) \\ &= L(t - t_i) + M(x_g(t) - x_{g_r}(t_i)) + M(t - t_i)(K + Lt_i + Mx_{g_r}(t_i)), \\ & \quad t_i \leq t \leq t_{i+1}. \end{aligned}$$

De (8) il résulte que la fonction  $u(t) = x_f(t) - x_{f_r}(t)$  satisfait à l'inégalité

$$|u'| \leq L(t - t_i) + M|u| + M(t - t_i)(K + Lt_i + M|x_{f_r}(t_i)|), \quad t_i \leq t \leq t_{i+1}$$

et (9) signifie que la fonction  $v(t) = x_g(t) - x_{g_r}(t)$  est la solution de l'équation

$$v' = L(t - t_i) + Mv + M(t - t_i)(K + Lt_i + Mx_{g_r}(t_i)), \quad t_i \leq t \leq t_{i+1}.$$

En vertu de l'hypothèse  $|x_{f_r}(t_i)| \leq x_{g_r}(t_i)$  et  $|u(t_i)| \leq v(t_i)$  il en vient que dans l'intervalle  $\langle t_i, t_{i+1} \rangle$  on a  $|u(t)| \leq v(t)$ .

De (4) et (5) il vient aussi

$$\begin{aligned} |x'_{f_r}(t)| &= |f(t_i, x_{f_r}(t_i))| \leq K + Lt_i + M|x_{f_r}(t_i)| \\ &\leq K + Lt_i + Mx_{g_r}(t_i) = x'_{g_r}(t_i), \quad t_i \leq t \leq t_{i+1}, \end{aligned}$$

ce qui; en vertu de l'hypothèse  $|x_{f_r}(t_i)| \leq x_{g_r}(t_i)$ , démontre que l'on a  $|x_{f_r}(t)| \leq x_{g_r}(t)$  dans l'intervalle  $\langle t_i, t_{i+1} \rangle$ .

Donc, de l'hypothèse que les inégalités (6) sont satisfaites pour  $t = t_i$  il vient qu'elles restent valables pour  $t_i \leq t \leq t_{i+1}$ . Comme les inégalités (6) sont évidemment vraies pour  $t = t_0 = 0$ , il en résulte qu'elles sont satisfaites dans tout l'intervalle  $\langle 0, a \rangle$  ce qui achève la démonstration du théorème.

4. Dans ce qui précède nous avons supposé que les nombres dérivés  $f_t(t, x)$  et  $f_x(t, x)$  du second membre de l'équation (1) sont bornés. Une condition de ce type est indispensable. Par exemple il ne suffit pas d'admettre que les fonctions  $f$  sont bornées.

Désignons par  $H$  la famille de fonctions  $f$  telles que 1° Toute fonction  $f \in H$  est continue dans l'ensemble  $\pi$  et satisfait à la condition

$$|f(t, x)| \leq 1.$$

2° Pour tout  $f \in H$  il existe une et seulement une solution de l'équation (1) satisfaisant à la condition (2).

On peut démontrer que pour la famille  $H$  la fonction

$$\sup_{\substack{0 \leq t \leq a \\ f \in H}} |x_f(t) - x_{f\tau}(t)|$$

ne dépend pas de  $\tau$  et pour tout  $\tau$  est égale à  $2a$ .