DEMONSTRATIO MATHEMATICA
Vol. VIII No 3 1975

Stanisław K. Zaremba

L'ERREUR DANS LE CALCUL DES INTÉGRALES DOUBLES PAR LA METHODE DES BONS TREILLIS

§ 1. Introduction

Rappelons quelques notations fréquemment employées quand il s'agit de bons treillis: Les coordonnées d'un point dans l'espace euclidien à s dimensions seront désignées par une lettre avec un indice inférieur allant de 1 à s; la même lettre soulignée désignera le point lui-même. Le pavé

$$0 \leqslant x_i \leqslant 1 \qquad (i = 1, \ldots, s)$$

sera désigné par Q^s. Pour un point arbitraire $\underline{h}$ on pose

$$R(\underline{h}) = \max(1, |h_1|) \ldots \max(1, |h_s|).$$

Si une fonction réelle $f(x)$ satisfait des conditions de régularité et d'extension par périodicité que nous n'avons pas besoin de rappeler ici (pour ces conditions et pour les transformations de l'intégrale permettant de satisfaire les conditions de périodicité, voir p. ex. [7]), cette fonction peut être développée en série multiple de Fourier

$$(1.1) \qquad f(\underline{x}) = \sum_{\underline{h} \in Z^s} c_{\underline{h}} \exp(2\pi i\, \underline{h} \cdot \underline{x}),$$

où le point désigne un produit scalaire et où les coefficients $c_{\underline{h}}$ satisfont

Cette publication a été rendue possible par l'octroi no. RD-65 do Conseil National de Recherche du Canada et par l'octroi du Ministère de l'Education du Québec (action concertée).

(1.2) $$\left|c_h\right| \leqslant K_r\, R(\underline{h})^{-r},$$

K_r étant une constante qui ne dépend que de f et r étant un entier positif; quand $r \geqslant 2$, la série est donc uniformément et absolument convergente.

Pour obtenir une approximation de l'intégrale

$$\int_{Q^s} f(\underline{x})\, d\underline{x},$$

on choisit un entier positif m et un vecteur $\underline{g}$ à coordonnées entières. On désigne alors par $p(g)$ le minimum de $R(\underline{h})$, quand $\underline{h} \neq \underline{0} = \langle 0,\ldots,0\rangle$ et

(1.3) $$\underline{g}\cdot\underline{h} \equiv 0 \pmod m.$$

On regarde l'expression

$$\frac{1}{m}\sum_{k=0}^{m-1} f\left(\frac{k}{m}\,\underline{g}\right)$$

comme une valeur approchée de l'intégrale; les coordonnées de $m^{-1}\,k\underline{g}$ devraient être réduites modulo 1, mais si l'on pense à f comme étant donnée par (1.1), cela n'est plus nécessaire. On voit facilement (p. ex.[4], [7]) que si $c_{\underline{h}} = 0$ dès que $R(\underline{h}) \geqslant p(\underline{g})$, la formule d'intégration est exacte.

Dans le cas général, on trouve ([2], [7]) que l'erreur d'intégration est bornée par $K_r\, P^{(r)}(\underline{g})$, où

$$P^{(r)}(\underline{g}) = \sum R(\underline{h})^{-r},$$

la somme se rapportant à tous les $\underline{h} \in Z^s$ différents de $\underline{0}$ et satisfaisant (1.3).

On sait que [7]

(1.4) $$P^{(r)}(\underline{g}) \leqslant \frac{2^{3s+1}(\log m)^{s-1}}{(r-1)!\,(\log 2)^{s-1}p(\underline{g})^{r}},$$

mais cette borne supérieure est fortement exagérée. Comme la recherche des bons treillis se réduisait à celle des $\underline{g}$ donnant la plus grande valeur possible de $p(\underline{g})$ ([3],[5], [7]), il est important d'obtenir une borne supérieure plus précise de $P^{(r)}(\underline{g})$ en fonction de $p(\underline{g})$ et plus particulièrement une borne de $P^{(2)}(\underline{g})$, car

$$P^{(r)}(\underline{g}) < p(\underline{g})^{2-r} P^{(2)}(\underline{g}) \qquad (r > 2)$$

et cette dernière borne supérieure paraît relativement satisfaisante quand on substitue à $P^{(r)}(\underline{g})$ une borne supérieure satisfaisante de cette quantité.

Le problème d'obtenir une borne supérieure de $P^{(2)}(\underline{g})$ meilleure que (1.4) dans plus de deux dimensions présente des difficultés sérieuses. Le cas de $s = 2$ est beaucoup plus simple grâce à la possibilité d'employer directement l'algorithme des fractions continues. Dans ce cas-là (mais seulement dans ce cas) l'exposant de $\log m$ dans (1.4) est le meilleur possible; cependant nous allons réduire considérablement la constante en nous servant d'un raisonnement complètement différent et plus direct que celui qui avait donné (1.4). Grâce à une étude un peu plus approfondie de l'ensemble des solutions de (1.3), on obtient dans les cas les plus importants des bornes inférieures de $P^{(2)}(\underline{g})$ et l'on arrive au résultat suivant dont l'application à la pratique est évidente.

On sait [6] que dans deux dimensions le rapport $p(\underline{g}) : m$ est le plus grand possible quand, en désignant par $\langle u_n \rangle$ la suite des nombres de Fibonacci $\left(u_1 = u_2 = 1 ; \ u_{n+2} = u_{n+1} + u_n \ (n = 1, 2, \ldots)\right)$ one pose

$$(1.5) \qquad n = u_n, \quad \underline{g} = \langle 1, \ u_{n-1} \rangle ,$$

ce qui donne $p(\underline{g}) = u_{n-2}$, ou bien

$$(1.6) \qquad m = 2u_n, \quad \underline{g} = \langle 1, \ 2u_{n-1} \rangle$$

ce qui donne $p(\underline{g}) = 2u_{n-2}$.

D'ailleurs, dans deux dimensions, on peut toujours poser

$$P^{(2)}(\underline{g}) = m^{-2} c \log m; \tag{1.7}$$

il se trouve que le coefficient c a une borne inférieure positive universelle et une borne supérieure finie pour certaines classes de treillis admettant des valeurs du module m arbitrairement grandes. Il est donc naturel de regarder c comme une mesure du rendement d'un treillis. Jusqu'à présent, les treillis définis par (1.6) n'avaient guère attiré l'attention, mais on verra que, jugé par ce critère, leur rendement est supérieur à celui des treillis définis par (1.5).

§ 2. Une majorante de $P^{(2)}(g)$

Nous nous bornons au cas où le nombre d'éléments du développement de g_2/m (avec $g_1 = 1$) n'est pas inférieur à 3. Ceci n'est pas une restriction sérieuse, car si ce nombre est inférieur à 3, m est très petit à moins qu'au moins un élément de cette fraction continue ne soit relativement grand, auquel cas [6] $p(g) : m$ est petit, ce qui rend aussi le treillis peu intéressant. D'ailleurs, il n'y aurait aucune difficulté à adapter le raisonnement qui suit au cas d'une fraction continue composée d'un ou deux éléments.

P r o p o s i t i o n 2.1. Soit $\underline{g} = \langle 1, g_2 \rangle$ avec $(g_2, m) = 1$, le développement de g_2/m en fraction continue comportant au moins trois éléments. Alors

$$P^{(2)}(\underline{g}) < \frac{\pi^4 \log n}{12\, p(\underline{g})^2 \log \xi}, \tag{2.01}$$

où $\xi = \frac{1}{2}(1 + \sqrt{5})$.

D é m o n s t r a t i o n. Désignons par $b_1, \dots, b_q$ les éléments du développement de g_2/m en fraction continue et par A_k/B_k ses réduites successives; en particulier $B_q = m$. Soit

$$B_{k-1} \leqslant h_2' < B_k; \quad B_{k-1} \leqslant h_2'' < B_k \qquad (k \geqslant 2) \tag{2.02}$$

On a alors $|h_2' - h_2''| < B_k$ et d'après le théorème de Lagrange sur les fractions continues, pour toute paire d'entiers t' et t'',

$$\left|(h_2' - h_2'')g_2 - (t' - t'')m\right| > \left|B_{k-1}\, g_2 - A_{k-1}\, m\right|.$$

Donc si $h_1^{(k-1)} = -B_{k-1}\, g_2 + A_{k-1}\, m$, $|h_1^{(k-1)}|$ est une borne inférieure des valeurs absolues des différences entre les valeurs de h_1, quand $\underline{g}$ satisfait (1.3) et h_2' et h_2'' satisfont (2.02). En tenant compte des quatre combinaisons possibles des signes de h_1 et h_2 on trouve, quand $k > 1$,

$$(2.03) \qquad \sum_{\substack{B_{k-1} \leqslant |h_2| < B_k \\ \underline{g}\cdot\underline{h} \equiv 0 \pmod{m}}} \frac{1}{R(\underline{h})^2} \leqslant \frac{4}{B_{k-1}^2\, h_1^{(k-1)^2}} \sum_{k=1}^{\infty} \frac{1}{k^2} \leqslant \frac{2\pi^2}{3\, p(\underline{g})^2}.$$

Soit maintenant $P_k^{(2)}(\underline{g})$ la somme de $R(\underline{h})^{-2}$ pour tous les $\underline{h}$ satisfaisant (1.3) et tels que h_2 soit congruent modulo m à un nombre h pour lequel $h_2\, h \geqslant 0$ et $B_{k-1} \leqslant |h| < B_k$. Quand $k \geqslant 2$, on obtient cette somme à partir de (2.03) en substituant dans chaque terme de cette dernière somme l'expression

$$\sum_{\lambda=0}^{\infty} \frac{1}{(|h_2| + \lambda m)^2}$$

à h_2^{-2}. Le rapport de cette somme à h_2^{-2} est

$$(2.04) \qquad \sum_{\lambda=0}^{\infty} \frac{1}{(1 + \lambda m/|h_2|)^2}.$$

Si l'on écrit la fraction continue de façon à éviter que le dernier élément soit égal à 1, on a $B_k < m/2$, quand $k \leqslant q-1$, et (2.04) est plus petite que

$$\sum_{\lambda=0}^{\infty} \frac{1}{(1+2\lambda)^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{3}{4} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{8}.$$

Si $k = q$, on se contente de la majorante

$$\sum_{\lambda=1}^{\infty} \frac{1}{(1+\lambda)^2} = \frac{\pi^2}{6}.$$

On a donc, d'après (2.03),

$$P_k^{(2)}(\underline{g}) \leqslant \frac{\pi^4}{12\, p(\underline{g})^2} \qquad (k = 2,\ldots,q-1) \tag{2.05}$$

et

$$P_q^{(2)}(\underline{g}) \leqslant \frac{\pi^4}{9\cdot p(\underline{g})^2}. \tag{2.06}$$

Quand à $P_1^{(2)}(\underline{g})$, on remarque que ni $p(\underline{g})$ ni $P^{(2)}(\underline{g})$ ne changent pas, quand on substitue $m - g_2$ à g_2. On peut donc supposer $g_2 > m/2$, ce qui entraîne $B_1 = 1$. Alors $P^{(2)}(\underline{g})$ est la somme de $R(\underline{h})^{-2}$ par rapport à tous les $\underline{h}$ satisfaisant $h_1 \equiv h_2 \equiv 0 \pmod{m}$ et $|h_1| + |h_2| > 0$. On trouve donc

$$P_1^{(2)}(\underline{g}) = \frac{4}{m^2} \sum_{k=1}^{\infty} \frac{1}{k^2} + \frac{4}{m^4}\left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right)^2 = \frac{2\pi^2}{3m^2} + \frac{\pi^4}{9m^4}. \tag{2.07}$$

D'après (2.05), (2.06) et (2.07), il vient

$$P^{(2)}(\underline{g}) < \frac{(q-2/3)\pi^2}{12\, p(\underline{g})^2} + \frac{2\pi^2}{3m^2} + \frac{\pi^4}{9m^4}. \tag{2.08}$$

Il n'y a plus qu'à trouver une bonne majorante de q en fonction de m. Evidemment, quand q est donné, m est le plus petit si tous les éléments de la fraction continue g_2/m sauf le dernier sont égaux à 1, celui-ci étant ègal à 2. En tenant compte du fait que la première réduite est alors $1/u_2$ (où $\langle u_k \rangle$ désigne toujours la suite des nombres de Fibonacci) et que si la dernière réduite a pour dénominateur u_n, celui de l'avant-dernière est u_{n-2}, on trouve

(2.09) $$m = B_q \geqslant u_{q+2}.$$

D'ailleurs, pour tout n

(2.10) $$u_n = 5^{-\frac{1}{2}}\left(\xi^n - (-\xi)^{-n}\right) = 5^{-\frac{1}{2}}\,\xi^n\left(1 + (-1)^{n+1}\,\xi^{-2n}\right),$$

où, comme précédemment, $\xi = \frac{1}{2}(1 + \sqrt{5})$.
Par suite

$$\log u_n = n \log \xi - \log \sqrt{5} + \log \left(1 + (-1)^{n+1}\,\xi^{-2n}\right),$$

donc

$$n = \left(\log u_n + \log \sqrt{5} - \log \left(1 + (-1)^{n+1}\,\xi^{-2n}\right)\right)/\log \xi \leqslant$$

$$\leqslant (\log u_n + \log \sqrt{5} - \log (1 - \xi^{-2n}))/\log \xi$$

et comme

$$\log(1-\alpha) > -\alpha/(1-\alpha) \quad (\alpha > 0),$$

on trouve finalement

(2.11) $$n < \left(\log u_n + \log \sqrt{5} + \xi^{-2n}\,(1 - \xi^{-2n})^{-1}\right)/\log \xi .$$

Puisque dans nos hypothèses $n \geqslant 5$, on a

$$\xi^{-2n}(1 - \xi^{-2n})^{-1} < 0.01$$

et, puisque $\log \sqrt{5} < 0.81$, on en déduit

$$q + 2 < \frac{\log m}{\log \xi} + \frac{0.82}{\log \xi} < \frac{\log m}{\log \xi} + \frac{41}{24}$$

et

$$q - \frac{2}{3} < \frac{\log m}{\log \xi} - \frac{23}{24}.$$

D'après (2.08), il vient

$$P^{(2)}(\underline{g}) < \frac{\pi^4 \log m}{12\, p(\underline{g})^2 \log \xi} - \frac{23\,\pi^4}{288\, p(\underline{g})^2} + \frac{2\pi^2}{3m^2} + \frac{\pi^4}{9m^4}. \tag{2.12}$$

D'ailleurs, on a toujours $p(\underline{g}) < (2/5)m$. On voit alors que la somme des deux derniers termes est plus petite en valeur absolue que le second. On a donc à plus forte raison l'inégalité (2.01).

En tenant compte des valeurs numériques de π et de $\log \xi$, on déduit de (2.01) l'inégalité suivante

$$P^{(2)}(\underline{g}) < 17\, p(\underline{g})^{-2} \log m; \tag{2.13}$$

en comparant ceci avec la majorante (1.4) de $P^{(2)}(\underline{g})$ précédemment trouvée, on voit que le coefficient de $p(\underline{g})^{-2} \log m$ a été réduit par un facteur de plus de 10.

§ 3. Une classe remarquable de treillis

Comme nous l'avions remarqué au début, les treillis définis par (1.5) méritent tout particulièrement notre attention. Dans ce cas-là

$$p(\underline{g})/m = u_{q-2}/u_q \geqslant 3/8 \tag{3.1}$$

ce qui nous permet d'exprimer la majorante (2.13) en fonction de m; il vient

$$P^{(2)}(\underline{g}) < 120 m^{-2} \log m. \tag{3.2}$$

Cependant, il nous a semblé utile d'examiner une classe de treillis qui généralise celle dont il vient d'être question. On définit la suite $\langle v_k \rangle$ par

$$v_0 = 0; \quad v_1 = 1; \quad v_{k+2} = bv_{k+1} + v_k \quad (k = 0,1,2,\dots),$$

où b est un entier positif. On pose ensuite

(3.3) $$m = v_q, \quad \underline{g} = \langle 1, v_{q-1} \rangle .$$

On remarque que, quand $b = 1$, on a les nombres de Fibonacci et les treillis définis par (1.5).

P r o p o s i t i o n 3.1. Pour le treillis défini par (3.3),

(3.4) $$p(g) = b\, v_{q-2} .$$

D é m o n s t r a t i o n. Ceci est une généralisation de la propriété des treillis engendrés par les nombres de Fibonacci qui était citée dans l'introduction. Le raisonnement qui suit est une simple généralisation de la démonstration de la Proposition 2.3 dans [6]. De la même façon on montre que, si $v_{k+1} \leqslant |h_2| \leqslant v_{k+2}$, (1.3) entraîne $R(\underline{h}) \geqslant b\, v_q\, v_{k+1}/v_{k+3}$ et l'on remarque que, les rôles de h_1 et h_2 étant symétriques, on peut supposer

(3.5) $$v_{k+2}^2 > h_2^2 \geqslant b\, v_q\, v_{k+1}/v_{k+3} .$$

Dans ce cas, le plus petit entier égal ou plus grand que cette borne inférieure $b\, v_q\, v_{k+1}/v_{k+3}$ de $R(\underline{h})$ est au moins $b\, v_{q-2}$. En effet,

$$\left| \frac{b\, v_q\, v_{k+1}}{v_{k+3}} - b\, v_{q-2} \right| = b\, v_q \left| \frac{v_{k+1}}{v_{k+3}} - \frac{v_{q-2}}{v_q} \right| \leqslant \frac{b^2\, v_q}{v_{k+3}\, v_{k+4}},$$

la dernière inégalité étant une conséquence facile des propriétés bien connues des fractions continues. D'ailleurs,

$$(b^2 + 1)/b > v_{k+3}/v_{k+2} > (b^2 + b + 1)/(b + 1).$$

Donc, en tenant compte de (3.5), on trouve

$$\left|\frac{b\ v_q\ v_{k+1}}{v_{k+3}} - b\ v_{q-2}\right| < \frac{b^2\ v_q(b+1)^3}{(b^2+b+1)^3\ v_{k+2}^2} \leqslant \frac{b(b+1)^3\ v_{k+3}}{(b^2+b+1)^3\ v_{k+1}} <$$

$$< \frac{(b+1)^3\ (b^2+1)^2}{(b^2+b+1)^3\ b} < 1$$

dès que $b \geqslant 2$, tandis que le cas de $b=1$ avait été traité précédemment. Donc $p(\underline{g}) \geqslant b\ v_{q-2}$, mais on a l'égalité, car $h_1 = v_{q-2}$, $h_2 = b$ satisfait (1.3).

Evidemment, les treillis du type considéré correspondant à des valeurs relativement grandes de b n'ont aucune importance au point de vue de l'analyse numérique, puisque [6] $p(\underline{g}) \leqslant m/b$. Cependant, il vaut peut-être la peine d'examiner d'un peu plus près le cas de $b = 2$, dans lequel le rapport $p(q):m$ n'est que légèrement plus petit que dans le cas des treillis construits à partir des nombres de Fibonacci, tandis que les valeurs de q en rapport avec m sont nettement plus petites. A cause de cela, on trouve la proposition suivante.

P r o p o s i t i o n 3.2. Quand $\underline{g}$ est déterminé par (3.3) avec $b = 2$ et $q \geqslant 4$,

$$P^{(2)}(\underline{g}) < \frac{\pi^4(\log m + 1{,}05)}{12\ q(\underline{g})^2\ \log \eta} - \frac{\pi^4}{18\ q(\underline{g})^2} + \frac{2\pi^2}{3m^2} + \frac{\pi^4}{9m^4},$$

où $\eta = 1 + \sqrt{2}$.

D é m o n s t r a t i o n. Comme dans le cas général, on a (2.08), mais en cherchant la valeur de q, on doit se souvenir que si, pour faire $g_2 > m/2$, on substitue $v_q - v_{q-1}$ à $v_{q-1} = g_2$, on allonge le développement en fraction continue de g_2/m, de sorte que $B_q = v_q$, mais on a toujours $B_{q-1} = v_{q-1} < m/2$. Pour trouver une majorante de q, qui est presque égale à sa valeur, on remarque que

$$v_k = 2^{-3/2}\left(\eta^k + (-1)^{k+1}\eta^{-k}\right) = 2^{-3/2}\eta^k\left(1 + (-1)^{k+1}\eta^{-2k}\right)$$

et que par suite

$$q = \left(\log v_q + \log (2^{3/2}) - \log\left(1 + (-1)^{k+1}\,\eta^{-2k}\right)\right)/\log \eta <$$

$$< \left(\log m + \log (2^{3/2}) + \eta^{-2k}\,(1 - \eta^{-2k})^{-1}\right)/\log \eta <$$

$$< \frac{\log m + 1.05}{\log \eta}$$

dès que $q \geqslant 4$. En substituant cette majorante de q dans (2.08), on trouve (3.6).

Comme dans tous les cas $p(\underline{g})/m \geqslant 1/3$, en substituant les valeurs numériques de π et $\log \eta$ dans (3.6), on trouve

$$P^{(2)}(\underline{g}) < 84m^{-2} \log m. \tag{3.7}$$

Ceci paraît beaucoup plus avantageux que (3.2), mais ce n'est qu'une comparaison de majorantes; en fait, les résultats des calculs numériques reproduits à la fin montrent que tout de même les treillis formés à partir des nombres de Fibonacci sont un peu plus avantageux, au moins quand m est dans l'intervalle pour lequel on a fait les calculs.

§ 4. Minorantes de $P^{(2)}(g)$

Il est difficile d'obtenir une bonne minorante de $P^{(2)}(\underline{g})$, quand seulement m et $p(\underline{g})$ sont donnés; cependant, on obient sans difficulté la proposition suivante

P r o p o s i t i o n 4.1. Pour tout treillis $g = \langle 1, g_2\rangle$ avec $(g_2, m) = 1$,

$$P^{(2)}(\underline{g}) > 4m^{-2} \log m. \tag{4.1}$$

D é m o n s t r a t i o n. En nous servant toujours des notations introduites précédemment et en supposant $g_2 > m/2$, nous nous bornons aux termes correspondant à $h_2 = \pm B_k$ ($k = 1,\dots,q-1$). Comme

$$\left| A_k/B_k - g_2/m \right| \leqslant 1/(B_k\ B_{k+1}) ,$$

il vient $|h_1| \leqslant m/B_{k+1}$ et $R(\underline{h}) \leqslant m\ B_k/B_{k+1}$, donc

$$R(\underline{h})^{-2} \geqslant m^{-2}\ (B_{k+1}/B_k)^2 > 2m^{-2} \log\ (B_{k+1}/B_k).$$

En faisant la somme par rapport à k et en tenant compte des deux signes possibles de h_2, on trouve (4.1).

On croit qu'une minorante obtenue pour un cas spécial que nous allons traiter est valable pour un treillis arbitraire. Cependant, en conjonction avec (3.2), la dernière proposition confirme d'une façon simple et élémentaire que, comme nous l'avions indiqué dans l'introduction, $m^2 P^{(2)}(\underline{g})/\log m$ est une norme raisonnable du rendement d'un treillis.

P r o p o s i t i o n 4.2. Si $m = u_q \geqslant 5$ et $\underline{g} = \langle 1, u_{q-1} \rangle$ modulo m,

$$P^{(2)}(\underline{g}) > 17m^{-2} \log m. \tag{4.2}$$

D é m o n s t r a t i o n. On commence à se borner de nouveau aux termes correspondant à $|h_2| = B_{k-1} = u_k$, mais cette fois on choisit le développement de g/m, où tous les éléments sont égaux à 1. D'ailleurs, on vérifie facilement que

$$u_k\ u_{q-1} - u_{k-1}\ u_q = (-1)^{k+1}\ u_{q-k}\ (k=0,1,\dots,q)$$

et l'on obtient la valeur de h_1 satisfaisant (1.3) et ayant la plus petite valeur absolue si l'on pose

$$h_1 = (-1)^{k+1}\ u_{q-k} \operatorname{sgn} h_2, \tag{4.3}$$

de sorte que

$$R(\underline{h}) = u_k\ u_{q-k}. \tag{4.4}$$

On vérifie que

$$u_k\ u_{q-k} \leqslant 2u_{q-3} \qquad (k = 2,3,\dots,q-2) \tag{4.5}$$

en remarquant que d'après (2.10)

$$u_k\, u_{q-k} = \Big(\big(\xi^q + (-1)^q\, \xi^{-q} + (-1)^{k+1}\, (\xi^{q-2k} + (-1)^q\, \xi^{2k-q})\big)\Big)/5 .$$

Soit

$$f(k) = \xi^{q-2k} + (-1)^q\, \xi^{2k-q} ;$$

il vient

$$f'(k) = 2 \log \xi \big(-\xi^{q-2k} + (-1)^q\, \xi^{2k-q}\big).$$

Si q est impair, alors $f'(k) < 0$ et la suite $\langle u_k\, u_{q-k} \rangle$ restreinte aux valeurs impaires de k est décroissante. Comme $k \geqslant 2$, le maximum correspond à $k = 3$ avec $u_k\, u_{q-k} = 2u_{q-3}$. Mais $u_k\, u_{q-k}$ est invariant par rapport à la substitution de $q-k$ à k et k est pair, quand $q-k$ est impair. Le maximum est donc valable pour tout k entier entre 2 et $q-2$.

Si q est pair, toujours d'après (2.10),

$$u_k\, u_{q-k} \leqslant (\xi^q + \xi^{-q} + \xi^{q-2k} + \xi^{2k-q})/5$$

et la fonction de k dans le second membre est visiblement décroissante tant que $k \leqslant q/2$ et invariante par rapport à la substitution de $q - k$ à k; son maximum dans l'intervalle $[3,\, m - 3]$ est donc atteint à $k = 3$, mais alors $q - 3$ est impair, de sorte que l'inégalité ci-dessus devient une égalité. Quand à $k = 2$ ou $k = q - 2$, $u_2\, u_{q-2} = u_{q-2} \leqslant 2u_{q-3}$ $(q=5,6,\dots)$. L'inégalité (4.5) est donc satisfaite dans tous les cas, puisque $u_3 = 2$. D'après (4.3), il s'en suit que $R(\underline{h}) \leqslant 2u_{q-3}$, quand $|h_2| = u_k$ $(k = 2,3,\dots,q-2)$ et h_1 est donné par (4.3). Comme $u_q \geqslant 4u_{q-3}$ $(q = 5,6,\dots)$, il vient

$$R(\underline{h})^{-2} > 8u_q^{-2} . \tag{4.6}$$

On a $q - 3$ termes de ce genre. Une légère modification du raisonnement du § 3 donne une minorante de q. Au lieu de (2.11) on écrit

$$q \geqslant \left(\log u_q + \log\sqrt{5} - \log\left(1 + \xi^{-2q}\right)\right)/\log\xi \geqslant$$

$$\geqslant\left(\log u_q + \log\sqrt{5} - \xi^{-2q}\right)/\log\xi > (\log u_q + 0.75)/\log\xi$$

dès que $q \geqslant 5$. Donc

$$q - 3 > (\log m - 9/4)\log\xi.$$

La somme des termes (4.6) est donc plus grande que

$$(8 \log m - 18)/(m^2 \log\xi) > (16.6 \log m - 37.41)m^{-2}.$$

Il y a encore deux termes correspondant à $k = 1$ et $k = n-1$, pour lesquels $R(\underline{h}) = u_1 u_{q-1} = u_{q-1} < (2/3)m$. En tenant compte des deux signes possibles dans chacun de ces deux termes, on trouve

$$16/(9m^2) > 1.77m^{-2}.$$

En ajoutant ceci à la somme précédente, on arrive à

$$(16.6 \log m - 35.64)m^{-2}.$$

Mais chacun des vecteurs $\underline{h}$ donnant lieu à cette somme peut encore être doublé. Alors $R(\underline{h})^{-2}$ est divisé par 16 et l'addition de ces termes multiplie la somme par 17/16, donnant

$$(17.6 \log m - 37.82)m^{-2}. \tag{4.7}$$

Finalement, il y a les termes correspondant à $h_1 \equiv 0 \pmod m$ ou $h_2 \equiv 0 \pmod m$, mais non pas $h_1 = h_2 = 0$. D'après (2.07), leur somme est plus grande que

$$2\pi^{-2}/(3m^2) > 6.57m^{-2}.$$

Ajouté à (4.7) cela donne

$$(17.6 \log m - 31,25)m^{-2}. \tag{4.8}$$

Quand $q \geq 13$, (4.2) est une conséquence immédiate de cette minorante, tandis que pour $5 \leq q \leq 12$ on l'obtient des résultats numériques cités à la fin.

§ 5. Le cas où $(g_2, m) > 1$; une classe de treillis particulièrement avantageux

Si $(g_2, m) = d$, évidemment $p(\underline{g}) \leq m/d$. Comme nous savons obtenir $p(\underline{g}) > m/3$ pour des valeurs de m arbitrairement grandes, les cas où $d > 2$ offrent peu d'intérêt. Nous nous bornerons donc au cas où $d = 2$, mais notre raisonnement pourrait être facilement généralisé de façon à traiter les cas où $d > 2$.

P r o p o s i t i o n 5.1. Soit $g = \langle 1, g_2 \rangle$ modulo m avec $(g_2, m) = 1$. Si $\underline{\tilde{g}} = \langle 1, 2g_2 \rangle$ modulo $\tilde{m} = 2m$, on a

$$P^{(2)}(\underline{\tilde{g}}) = P^{(2)}(\underline{g})/4 + \pi^2 \tilde{m}^{-2}. \tag{5.1}$$

D é m o n s t r a t i o n. Si

$$\underline{\tilde{g}} \cdot \underline{\tilde{h}} \equiv 0 \pmod{\tilde{m}}, \tag{5.2}$$

$\tilde{h}_1$ est nécessairement pair. En posant $\tilde{h}_1 = 2h_1$, $\tilde{h}_2 = h_2$, on trouve que (5.2) est équivalent à (1.3). Ceci établit une correspondance biunivoque entre les solutions des deux congruences avec $R(\underline{\tilde{h}}) = 2R(\underline{h})$ à moins que $\tilde{h}_1 = 0$, auquel cas $R(\underline{\tilde{h}}) = R(\underline{h})$. On a donc

$$P^{(2)}(\tilde{g}) = \frac{1}{4} P^{(2)}(g) + r,$$

où r est égal aux trois-quarts de la somme des termes de $P^{(2)}(\underline{g})$ correspondant à $h_1 = 0$. Dans ces termes, $h_2 = km$ $(k = \pm 1, \pm 2, \ldots)$ et leur somme est

$$\frac{2}{m^2} \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{3m^2}.$$

Donc

$$r = (3/4)\pi^2/(3m^2) = \pi^2 \tilde{m}^{-2}.$$

Nous avions remarqué que si l'on définit c par (1.7), ce paramètre mesure, en un certain sens, le rendement du treillis. Parmi les treillis engendrés par des vecteurs $\underline{g}$ dont les deux coordonnées sont relativement premières par rapport à m, ceux qui sont engendrés par $\underline{g} = \langle 1, u_{q-1} \rangle$ modulo u_q ont vraisemblablement le meilleur rendement mesuré de cette façon. Si l'on passe au treillis engendré par $\underline{\tilde{g}} = \langle 1, 2u_{q-1} \rangle$ modulo $\tilde{m} = 2u_q$, on a d'après (5.1)

$$P^{(2)}(\underline{\tilde{g}}) = \tilde{m}^{-2} (c \log \tilde{m} - c \log 2 + \pi^2) < \tilde{m}^{-2} c \log \tilde{m},$$

puisque, d'après (4.2), $c > 17$ et $17 \log 2 > \pi^2$. Ceci montre que le rendement des treillis engendrés par $\underline{\tilde{g}}$ modulo $\tilde{m}$ est (légèrement) supérieur à celui des treillis engendrés par $\underline{g} = \langle 1, u_q \rangle$ modulo m. Un raisonnement semblable montrerait que le résultat de la comparaison serait le même si l'on se basait sur $P^{(n)}(\underline{g})$ $(n > 2)$ au lieu de $P^{(2)}(\underline{g})$. D'ailleurs $p(\underline{\tilde{g}}) : \tilde{m} = p(\underline{g}) : m$. On a donc toutes les raisons de préférer $\underline{\tilde{g}}$ modulo $\tilde{m}$ à $\underline{g}$ modulo m, du moins tant qu'il s'agit d'intégrandes développables en séries de Fourier; quand ce développement n'est pas convergent, d'autres ensembles de points, qui ne sont pas des treillis, peuvent être plus avantageux.

Revenant encore aux treillis, on remarque que les treillis où ni g_1 ni g_2 ne sont relativement premiers par rapport à m ne peuvent pas être avantageux. En effet, si $(g_1, g_2, m) = d$, le treillis est équivalent au treillis engendré par $d^{-1} \underline{g}$ modulo m/d. Si $(g_1, g_2, m) = 1$, mais g_1 et g_2 ont chacun un facteur commun plus grand que 1 avec m, au moins un de ces facteurs est nécessairement égal ou plus grand que 3, ce qui entraîne $p(\underline{g})/m < 1/3$.

§ 6. Résultats numériques

Les résultats ci-dessous ont été obtenus par J.M.St-Pierre avec l'aide de l'ordinateur IBM 370 de l'Université de Sherbrooke; quand au principe du calcul de $P^{(2)}(\underline{g})$, voir p. ex. [5] ou [7].

TABLE 1

$n = u_q \qquad g_1 = 1,\ g_2 = u_{q-1}$

m	g_2	$P^{(2)}(\underline{g})$	$m^2\, P^{(2)}(\underline{g})/\log m$
5	3	2,275	35,35
8	5	1,080	33,25
13	8	$4,759 \times 10^{-1}$	31,35
21	13	$2,091 \times 10^{-1}$	30,29
34	21	$8,975 \times 10^{-2}$	29,42
55	34	$3,815 \times 10^{-2}$	28,80
89	55	$1,603 \times 10^{-2}$	28,29
144	89	$6,685 \times 10^{-3}$	27,89
233	144	$2,767 \times 10^{-3}$	27,56
377	233	$1,139 \times 10^{-3}$	27,28
610	377	$4,662 \times 10^{-4}$	27,05
987	610	$1,900 \times 10^{-4}$	26,85
1597	987	$7,713 \times 10^{-5}$	26,67
2584	1597	$3,120 \times 10^{-5}$	26,51

TABLE 2

$m = v_q \qquad g_1 = 1,\ g_2 = v_{q-1}$

m	g_2	$P^{(2)}(\underline{g})$	$m^2\, P^{(2)}(\underline{g})/\log m$
12	5	$5,562 \times 10^{-1}$	32,23
29	12	$1,224 \times 10^{-1}$	30,58
70	29	$2,570 \times 10^{-2}$	29,64
169	70	$5,213 \times 10^{-3}$	29,03
408	169	$1,032 \times 10^{-3}$	28,59
985	408	$2,008 \times 10^{-4}$	28,26
2378	985	$3,851 \times 10^{-5}$	28,01
5741	2378	$7,304 \times 10^{-6}$	27,81

LITTERATURE

[1] E. H l a w k a: Funktionen von beschränkter Variation in der Theorie der Gleichverteilung, Ann.Mat. Pura Appl. (IV) 54 (1961) 325-334.

[2] E. H l a w k a: Zur angenäherten Berechnung mehrfacher Integrale, Monatsh. Math. 66 (1962) 140-151.

[3] G. K e d e m, S.K. Z a r e m b a: A table of good lattice points in three dimensions, à paraître dans Numerische Math.

[4] N.M. K o r o b o v: Teoretikocislovye metody v priblizhennom analize (Méthodes de la théorie des nombres dans l'analyse approchée), Fizmatgiz, Moscou 1963.

[5] D. M a i s o n n e u v e: Recherche et utilisation des "Bons Treillis". Programmation et résultats numériques, pp. 121-200, dans Applications de la Théorie des nombres à l'Analyse numérique, New York - London 1972.

[6] S.K. Z a r e m b a: Good lattice points, discrepancy, and numerical integration, Ann. Mat. Pura Appl. (iv) 73 (1966) 293-318.

[7] S.K. Z a r e m b a: La méthode des "Bons Treillis" pour le calcul des intégrales multiples, pp. 39-116 dans Applications de la Théorie des Nombres à l'Analyse Numérique, New York - London 1972.

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DE SHERBROOKE, SHERBROOKE, QUÉBEC, CANADA J1K 2R1

Received February 10, 1975.