Ann. Sc. Math. Québec, 1977, vol. I, No. 2, pp. 175-182

BASES D'UN NOUVEL ALGORITHME POUR LA DETERMINATION DES "BONS TREILLIS"

Marc Bourdeau

§1. Les bons treillis.

Soit h un vecteur de $\mathbb{Z}^n$, noté $h = (x_1,\dots,x_n)$. On définit R(h), le produit des valeurs absolues des coordonnées de h différentes de 0:

$$R(h) = \prod_{i=1}^{n} \operatorname{Max}(1,|x_i|).$$

Soit m, un entier positif (on l'appelle le module) et $g \in \mathbb{Z}^n$, avec $g_1 = 1$ et g_i, les diverses coordonnées, réduites modulo m. On pose

$$H(m;g) = \{h \in \mathbb{Z}^n : h \neq 0,\ g\cdot h \equiv 0 \pmod m\},$$

où $g\cdot h$ dénote le produit scalaire usuel de $\mathbb{R}^n$, et on définit

$$\rho(g) = \min_{H(m,g)} R(h).$$

On sait que, lorsque l'on veut approximer l'intégrale d'une fonction à n variables définies sur $[0,1]^n$ par

$$\frac{1}{m}\sum_{k=0}^{m-1} f(\tfrac{k}{m}\,g), \quad \text{avec les coord. de } \tfrac{k}{m}\,g \text{ réduites modulo 1,}$$

l'erreur que l'on commet possède des bornes d'autant plus favorables, sur certaines classes bien définies de fonctions, que $\rho(g)$ est grand. Le professeur Zaremba a d'ailleurs obtenu les meilleures bornes connues jusqu'à maintenant [1].

Le problème de trouver, pour m donné, $g_0 \in \mathbb{Z}^n$ (t.g. $g_1 = 1$, $0 \le g_i \le m-1$) pour lequel

$$\rho(g_0) \equiv \rho_0 = \max_g\{ \min_{H(m,g)} R(h)\},$$

n'est résolu, de façon satisfaisante pour les usages pratiques actuels, que pour $n \leq 4$ (n est le nombre de variables de la fonction à intégrer). Les méthodes actuelles proposent diverses façons de réduire l'ensemble G des g possibles, ainsi que l'ensemble $H(m;g)$ [2], [3]. Dans la programmation proprement dite(*), nous aurons à utiliser certaines réductions pour l'ensemble G déjà utilisées par Dominique Maisonneuve. Pour le moment, nous allons donner les fondements d'une méthode de réduction de $H(m,g)$ en utilisant certaines régularités de la fonction $R(\cdot)$. Cette méthode aura, par surcroît, l'avantage de réduire encore plus G.

§2. La méthode.

Dans ce qui suit, on admettra que $n \geq 3$, et que m est fixé.

La première chose qu'on remarque, en utilisant des symétries dans la définition de $R(\cdot)$, est qu'on peut restreindre notre recherche du minimum de $R(\cdot)$ sur $H(m,g)$ à l'ensemble $\tilde{H}(m,g)$ où

$$\tilde{H}(m;g) = \{h \in H(m;g) : |x_i| \leq m-1\}.$$

On pourrait même faire légèrement mieux ($|x_i| \leq \lfloor m/2 \rfloor$, [2]), mais cela ne nous serait d'aucune utilité par la suite. Dans ce qui suit on commettra l'abus de langage suivant: $H(m;g)$ sera en fait $\tilde{H}(m;g)$.

Une fois qu'on admet cette première restriction, pour un g donné, nous divisons l'examen des divers éléments de $H(m,g)$ en 2^{n-2} parties, correspondant chacune à un des "quadrants"

$$\begin{cases} x_2 \geq 0 & \text{ou } x_2 < 0 \\ x_3 \geq 0 & \text{ou } x_3 < 0 \\ \vdots & \\ x_{n-1} \geq 0 & \text{ou } x_{n-1} < 0, \end{cases}$$

de la façon suivante: donnons-nous $x_2,\ldots,x_{n-1}$ des coordonnées pour $h \in H(m,g)$. Pour chaque x_n, x_1 est déterminée par la congruence $h\cdot g \equiv 0(m)$.(**) On pose

(*) Les résultats numériques de notre méthode sont encore à venir.

(**) Dans ce qui suit (m) signifie $(\mathrm{mod}\ m)$. De plus un nombre entre parenthèses indicées par m, est le nombre mod m: $(x_i)_m$ est x_i mod m. Et $\lfloor x \rfloor$ signifie le plus grand entier dans x.

$$H(m;g;x_2,\dots,x_{n-1}) = \{h \in H(m,g) : x_2,\dots,x_{n-1} \text{ sont fixés};$$
$$x_n \text{ varie de } -(m-1) \text{ à } m-1,\ x_1 \text{ est déterminé par } h\cdot g \equiv 0(m)\}$$

et

$$H(m;g;\geq,\geq,\dots,\geq) = \{H(m;g;x_2,\dots,x_{n-1}) : 0 \leq x_i \leq m-1,\ i = 2,n-1\}.$$

Similairement, suivant le signe des x_i $(i = 2,\dots,n-1)$, on détermine les ensembles $H(m;g;\lesseqgtr,\dots,\lesseqgtr)$, où le sens de l'inégalité ($\leq$ ou $\geq$) est fixé par le signe de la coordonnée correspondante. On construit ainsi, pour m et g, une famille d'ensembles qui recouvrent $H(m;g)$, en le divisant en 2^{n-2} parties, correspondantes aux 2^{n-2} "quadrants", chacune étant notée par la bonne variante de $H(m;g;\gtreqless,\lesseqgtr,\dots,\lesseqgtr)$. Nous notons par $V(m;g;x_2,\dots,x_{n-1})$ l'ensemble des 2n-1 valeurs de $R(\cdot)$ pour les 2n-1 vecteurs de $H(m;g;x_2,\dots,x_{n-1})$, et on l'appelle le *fragment* de rang $x_2,\dots,x_{n-1}$. La signification des $V(m;g;\lesseqgtr,\dots,\lesseqgtr)$ est immédiate.

Dans ce qui suit, nous allons élaborer la base d'un algorithme qui étudie $H(m;g;\geq,\geq,\dots,\geq 0)$ le "premier" quadrant. Cependant nous voyons immédiatement comment ramener les autres à ce premier en modifiant g de façon convenable.

Par exemple, pour étudier le quadrant $H(m;g;\geq,\geq,\dots,\geq,\leq)$, on définit $\bar{g} = (1,g_2,\dots,m-g_{n-1},g_n)$. Soit $h \in H(m;g;\geq,\geq,\dots,\leq)$. On a $h = (x_1,\dots,x_n)$ où $x_{n-1} \leq 0$ et

$$x_1 + x_2g_2 + \dots + x_{n-1}g_{n-1} + x_ng_n \equiv 0(m).$$

Mais on a aussi:

$$x_1 + x_2g_2 + \dots + (-x_{n-1})(m-g_{n-1}) + x_ng_n \equiv 0(m),$$

de sorte que $h' = (x_1,x_2,\dots,-x_{n-1},x_n) \in H(m;\bar{g};\geq,\geq,\dots,\geq)$ et $R(h') = R(h)$. Comme on a aussi l'inverse, on a que

$$V(m;g;\geq,\geq,\dots,\geq,\leq) = V(m;\bar{g};\geq,\geq,\dots,\geq).$$

Plus généralement, pour étudier le minimum de $R(h)$ sur l'un ou l'autre des $H(m;g;\gtreqless,\dots,\gtreqless)$, on considère $\bar{g}$ où chacune des coordonnées de $\bar{g}$ est g_i ou $m-g_i$ suivant le sens de l'inégalité, et on examine le fragment $V(m;\bar{g};\geq,\dots,\geq)$.

Notre méthode consiste à examiner successivement les éléments d'une classe admissible pour g et à éliminer les candidats g pour lesquels $\rho(g) < \rho_0$, où ρ_0 est la meilleure borne déjà trouvée pour $\rho(\cdot)$. Elle a l'avantage que dès que g est éliminé, les 2^{n-2} $\bar{g}$ possibles le sont aussi. Prenons un exemple qui fera voir pourquoi.

Supposons que $g = (1,g_2,\dots,g_5)$ est éliminé, alors il existe $h \in H(m;g;\geq,\geq,\leq)$, disons, où $R(h) < \rho_0$; immédiatement, on a que $\bar{g} = (1,g_2,g_3,m-g_4,g_5)$ est éliminé car

$$V(m;\bar{g};\geq,\geq,\geq) = V(m;g;\geq,\geq,\leq),$$

et un des éléments de $V(m;g,\geq,\geq,\leq)$ est $< \rho_0$.

Si, disons $\bar{g}' = (1,g_2,m-g_3,g_4,g_5)$ n'est pas éliminé, alors, pour tous les quadrants $H(m;\bar{g}';\lesseqgtr,\lesseqgtr,\lesseqgtr)$, les éléments correspondants de $V(m;\bar{g}';\gtreqless,\lesseqgtr,\lesseqgtr)$ sont tous $\geq \rho_0$. En particulier, cela tient pour $V(m;\bar{g}';\geq,\geq,\leq)$. Or un raisonnement comme ci-haut nous donne que

$$V(m;\bar{g}';\geq,\geq,\leq) = V(m;\bar{g}'';\geq,\geq,\geq)$$

où

$$\begin{aligned}\bar{g}'' &= (1,\bar{g}_2',m-\bar{g}_3',m-\bar{g}_4',\bar{g}_5) \\ &= (1,g_2,g_3,m-g_4,g_5) \\ &= \bar{g}.\end{aligned}$$

Donc $V(m;\bar{g}';\geq,\leq,\leq) = V(m;\bar{g};\geq,\geq,\geq)$; mais un des éléments de ce dernier fragment est $< \rho_0$. C'est une contradiction.

Donc, dès que g est éliminé, les 2^{n-2} $\bar{g}$ possibles le sont aussi.

Concentrons-nous maintenant sur $H(m;g;\geq,\dots,\geq)$. Pour notre problème, nous devons en examiner les divers $V(m;g;x_2,\dots,x_{n-1})$ où $0 \leq x_i \leq m-1$, $i = 2,\dots,n-1$.

Nous allons démontrer que nous n'avons pas à examiner tous les fragments, c'est-à-dire tous les candidats possibles de $V(m;g;\geq,\dots,\geq)$, pour trouver le min de $R(\cdot)$ lorsque m et g sont donnés. Il y a, en effet, un certain nombre de $V(m;g;x_2,\dots,x_{n-1})$, appelés fragments initiaux (il y en a au plus m) pour la bonne raison que les éléments de tous les autres sont un multiple terme à terme des éléments de ces fragments initiaux. Nous avons pu estimer, par l'examen de 1000 divers m et g (*) que le facteur multiplicatif associé a un fragment est ≥ 1 pour 99% des fragments. De plus, cette proportion a tendance à augmenter lorsque m grandit. Nous avons un algorithme qui élimine *a priori* 50% (estimé grossier) des $V(m;g;x_2,\dots,x_{n-1})$ où le facteur multiplicatif est ≥ 1, et qui élimine d'un calcul simple, presque tous les autres. Naturellement, tous les fragments ayant un facteur multiplicatif < 1 sont examinés, et, de tous les fragments examinés, seulement 20% ont un facteur multiplicatif ≥ 1.

Nous espérons que dans la pratique notre algorithme sera plus efficace que ses prédécesseurs et permettra d'aborder l'étude des dimensions 5 et 6.

Voici donc comment se fait le passage d'un fragment à un fragment initial, ainsi que la détermination du facteur multiplicatif.

(*) En dimension 5 seulement.

Définition.

Soit $(1,g_2,\ldots,g_n)$ un candidat au titre de bon vecteur. Posons(*)

$$d_1 = m/(g_{n-1},m),$$

$$d_2 = (g_{n-1},m)/(g_{n-2},g_{n-1},m),$$

$$d_{n-2} = (g_3,g_4,\ldots,m)/(g_2,g_3,\ldots,g_{n-1},m).$$

Les fragments $V(m;g;x_2,\ldots,x_{n-1})$ où $0 \le x_{n-1} \le d_1-1,\ldots,0 \le x_2 \le d_{n-2}-1$ sont appelés les *fragments initiaux* de $V(m,g;\ge,\ldots,\ge)$.

Comme notre classe admissibles de g [2] a la propriété que $1 < g_2 < g_3 <\ldots< g_{n-1}$, les d_j sont bien définis. On remarque, de plus, qu'il y a exactement $m/(g_2,\ldots,g_{n-1},m)$ fragments initiaux.

D'où proviennent ces divers d_j: Pour j=1, d_1 est la plus petite solution de $g_{n-1}x_{n-1} \equiv 0(m)$. Si $d_1 = 1$, on a évidemment $d_j = 1$ pour $j > 1$. Pour $j > 1$, $k = d_j$ est le plus petit entier positif t.q.

$$g_{n-j}x_{n-j} +\ldots+ g_{n-1}x_{n-1} \equiv kg_{n-j-1}(m).$$

On est maintenant en mesure de démontrer le résultat clé de notre méthode.

Théorème.

On définit, pour m et g donnés, les d_j comme ci-haut. Si on écrit

$x_3^{(0)},\ldots,x_{n-1}^{(0)}$ pour la "plus petite solution"(**) de

$$g_3x_3 +\ldots+ g_{n-1}x_{n-1} \equiv d_{n-2}g_2(m) \tag{1}$$

$x_4^{(1)},\ldots,x_{n-1}^{(1)}$ pour la "plus petite solution" de

$$g_4x_4 +\ldots+ g_{n-1}x_{n-1} \equiv d_{n-3}g_3(m)$$

...

(*) Si $a_1,\ldots,a_k$ sont des naturels, $(a_1,\ldots,a_k)$ dénote leur p.g.c.d.

(**) "Plus petite solution" au sens où $x_3^{(0)}$ est le plus petit entier ≥ 0 t.q. (1) possède une solution, $x_4^{(0)}$ est le plus petit entier ≥ 0 t.q. (1) avec x_3 remplacé par $x_3^{(0)}$ possède une solution, etc...

$$x_{n-1}^{(n-4)} \text{ pour la plus petite solution de}$$

$$g_{n-1}x_{n-1} \equiv d_2 g_{n-2}(m),$$

on a, en écrivant x_2 sous la forme suivante: $x_2 = k_2 d_{n-2}+h_2$, $0 \le h_2 \le d_{n-2}-1$, que

$$V(m;g;k_2d_{n-2}+h_2,x_3,\ldots,x_{n-1}) =$$

$$\mathcal{L}\times V(m;g;h_2,(k_2x_3^{(0)}+x_3)_{d_{n-3}},(k_2x_4^{(0)}+k^{(1)}x_4^{(1)}+x_4)_{d_{n-4}},\ldots, \tag{2}$$

$$(k_2x_{n-1}^{(0)}+k^{(1)}x_{n-1}^{(1)}+\ldots+k_{n-1}^{(n-4)}x_{n-1}^{(n-4)}+x_{n-1})_{d_1}),$$

où on a posé

$$k^{(1)} = \left\lfloor \frac{k_2x_3^{(0)}+x_3}{d_{n-3}} \right\rfloor$$

$$k^{(2)} = \left\lfloor \frac{k_2x_4^{(0)}+k^{(1)}x_4^{(1)}+x_4}{d_{n-4}} \right\rfloor,\ldots,k^{(n-4)} = \left\lfloor \frac{k_2x_{n-2}^{(0)}+\ldots+k^{(n-5)}x_{n-2}^{(n-5)}+x_{n-2}}{d_2} \right\rfloor.$$

C'est-à-dire que tout fragment est un multiple d'un fragment initial, puisque tous les x_i dans l'ensemble de droite sont réduits modulo d_{n-i}. Bien sûr, le signe égal de (2) signifie que les éléments de l'ensemble à gauche sont un multiple $\mathcal{L}$ des éléments correspondants de l'ensemble à droite. Il reste à spécifier $\mathcal{L}$:

$$\mathcal{L} = \frac{\mathrm{Max}(k_2d_{n-2}+h_2,1)\mathrm{Max}(x_3,1)\ \ldots\ \mathrm{Max}(x_{n-1},1)}{\mathrm{Max}(h_2,1)\cdot\mathrm{Max}((k_2x_3^{(0)}+x_3)_{d_{n-3}},1)\ \ldots\mathrm{Max}((k_2x_{n-1}^{(0)}+\ldots+k^{(n-4)}x_{n-1}^{(n-4)}+x_{n-1})_{d_1},1)}.$$

<u>Démonstration</u>.

Donnons ici le premier élément de la preuve, très technique mais simple; les autres parties en sont la continuation naturelle.

Par définition de d_{n-2}, et $x_3^{(0)},\ldots,x_{n-1}^{(0)}$, on peut écrire un des éléments de

$$V(m;g;k_2d_{n-2}+h_2,0,0,\ldots,0),$$

suivant la valeur de x_n:

$$\mathrm{Max}(k_2d_{n-2}+h_2,1)\mathrm{Max}(|x_n|,1)\mathrm{Max}((-[k_2d_{n-2}+h_2]g_2-x_ng_n)_m,1) \tag{3}$$

puisque $x_1 \equiv (-x_2g_2-\ldots-x_ng_n)(\mathrm{mod}\ m)$ et $x_3 = x_4 =\ldots= x_{n-1} = 0$. L'élément correspondant de $V(m;g;h_2,(k_2x_3^{(0)})_m,\ldots,(k_2x_{n-1}^{(0)})_m)$ est

$$\begin{aligned}&\mathrm{Max}(h_2,1)\mathrm{Max}(|x_n|,1)\mathrm{Max}((k_2x_3^{(0)})_m,1)\ \ldots\ \mathrm{Max}((k_2x_{n-1}^{(0)})_m,1)\\ &\times\ \mathrm{Max}((-h_2g_2-k_2x_3^{(0)}g_3-\ldots-k_2x_{n-1}^{(0)}g_{n-1}-x_ng_n)_m,1).\end{aligned}\tag{4}$$

Et, par le choix même de $x_3^{(0)},\ldots,x_{n-1}^{(0)}$,

$$h_2g_2+k_2[x_3^{(0)}g_3+\ldots+x_{n-1}^{(0)}g_{n-1}]+x_ng_n\equiv(h_2g_2+k_2d_{n-2}g_2+x_ng_n)\ (\mathrm{mod}\ m)$$

ce qui fait que l'on peut écrire sommairement

$$\begin{aligned}V(m;g;k_2d_{n-2}+h_2,0,\ldots,0)&=\frac{\mathrm{Max}(k_2d_{n-2}+h_2,1)}{\mathrm{Max}(h_2,1)\mathrm{Max}((k_2x_3^{(0)})_m,1)\ldots\ \mathrm{Max}((k_2x_{n-1}^{(0)})_m,1)}\\ &\times\ V(m;g;h_2,(k_2x_3^{(0)})_m,\ldots,(k_2x_{n-1}^{(0)})_m)\end{aligned}\tag{5}$$

ou encore

$$V(m;g;k_2d_{n-2}+h_2,x_3,\ldots,x_{n-1})=\mathcal{L}_1\times V(m;g;h_2,(k_2x_3^{(0)}+x_3)_m,\ldots,(k_2x_{n-1}^{(0)}+x_{n-1})_m).$$

Maintenant on peut utiliser la définition de $x_4^{(1)},\ldots,x_{n-1}^{(1)}$ pour ramener $V(m;g;h_2,x_3,\ldots,x_{n-1})$ à

$$V\left(m;g;h_2,(x_3)_{d_{n-3}},\left(\left\lfloor\frac{x_3}{d_{n-3}}\right\rfloor x_4^{(1)}+x_4\right)_m,\ldots,\left(\left\lfloor\frac{x_3}{d_{n-3}}\right\rfloor x_{n-1}^{(1)}+x_{n-1}\right)_m\right).$$

Etc... A la fin, on obtient (2) où $\mathcal{L}$ est le produit des $\mathcal{L}_i$.

Puisque tous les fragments sont des multiples des fragments initiaux, notre méthode vise à éliminer, *a priori* ou par calcul, le plus possible des fragments qui se ramènent à un fragment initial par un facteur multiplicatif ≥ 1.

Notre algorithme repose sur le théorème précédent ainsi que sur le comportement, suivant les valeurs de x_0 et d, de la fonction:

$$f(x)=\frac{\mathrm{Max}(x,1)}{\mathrm{Max}((x_0+x)_d,1)},\qquad x\in\mathbb{N}.$$

Bibliographie

[1] S.K. Zaremba, La méthode des bons treillis pour le calcul des intégrales multiples, in *Applications of Number Theory to Numerical Analysis*, S.K. Zaremba, éditeur, Academic Press, New York, 1972, p. 39-119.

[2] D. Maisonneuve, Recherche et utilisation des "bons treillis". Programmation et résultats numériques, in *Applications of Number Theory to Numerical Analysis*, S.K. Zaremba, éditeur, Academic Press, New York, 1972, p. 121-201.

[3] G. Kedem, The Search for Good Lattice Point in N Dimension. University of Wisconsin, Mathematical Research Center, Technical Report No. 1570, October 1975.

Département de Mathématiques
Ecole Polytechnique
C.P.6079, Succursale A
Montréal, Québec, Canada H3C 3A7