Concours SMF Junior 2017

Rapports détaillés des auteurs/correcteurs et solutions

Problème 1 : Algèbre

Auteur : Reinhard Schäfke (Université de Strasbourg)

Soient $p, q \geq 2$ deux entiers multiplicativement indépendants, c'est-à-dire tels que, pour tout couple (a, b) d'entiers strictement positifs, on a $p^a \neq q^b$. On suppose que $f(x)$ et $g(x)$ sont deux polynômes complexes sans terme constant tels que

$$f(x^q) - f(x) = g(x^p) - g(x). \tag{1}$$

Montrer qu'il existe un polynôme complexe $h(x)$ tel que

$$h(x^p) - h(x) = f(x) \text{ et } h(x^q) - h(x) = g(x). \tag{2}$$

Rapport

Une équation contenant $h(x)$ et ses images successives par l'opérateur σ_p, $\sigma_p(h(x)) = h(x^p)$, est appelée "équation de p-Mahler". Un exemple simple est l'équation additive $h(x^p) - h(x) = f(x)$ avec un polynôme complexe f sans terme constant. En général, il n'en existe pas de solution polynomiale. Une condition nécessaire simple est $f(1) = 0$.

Le sujet pose la question de l'existence d'une solution polynomiale commune d'un système (2) d'équations additives de p- et de q-Mahler avec des polynômes f, g. Le calcul $h(x^{pq}) = \sigma_p \sigma_q(h(x))$ de deux manières montre que la condition (1) s'avère nécessaire. On l'appelle donc *condition de consistence* du système (2).

Le sujet demande de montrer que la condition (1) est aussi suffisante lorsque p et q sont multiplicativement indépendants. La condition sur p, q est naturelle et nécessaire : si $p^a = q^b$, alors $f(x) = x + x^q + \ldots + x^{q^{b-1}}$ et $g(x) = x + x^p + \ldots + x^{p^{a-1}}$ satisfont (1), mais il n'existe pas de polynôme h avec (2) car $f(1)$ et $g(1)$ sont non nuls.

Le sujet est un cas particulier du corollaire 22 de la prépublication [Schae-Si]. Une version simplifiée de ce corollaire est la suivante : un système

$$Y(x^p) = A(x)Y(x), \quad Y(x^q) = B(x)Y(x) \tag{3}$$

avec des matrices carrées $A(x), B(x)$ à coefficients polynomiaux telles que $A(0) = B(0) = I$ et $\det A(x) = \det B(x) \equiv 1$ admet une solution polynomiale Y avec $Y(0) = I$ et $\det Y(x) \equiv 1$, si p et q sont multiplicativement indépendants et si A et B satisfont la condition de consistence

$$A(x^q)B(x) = B(x^p)A(x). \tag{4}$$

Le sujet en devient un cas particulier si on pose $A(x) = \begin{pmatrix} 1 & 0 \\ f(x) & 1 \end{pmatrix}$.

35 équipes ont proposé une solution du sujet. Pour la correction, l'aide d'Augustin Fruchard (UHA Mulhouse) m'a été précieuse. Dans 21 de ces propositions, on trouve une solution essentiellement complète. Le sujet n'étant pas trivial, c'est un très bon résultat. Encore plus satisfaisant, presque étonnant, est la diversité et l'originalité des approches utilisées, même si la plupart des solutions suivent avec quelques variations la piste de la solution officielle : établir l'existence d'une solution h série formelle et ensuite montrer qu'elle est en fait polynomiale. Une solution ramène le sujet à une question de cohomologie – malheureusement sans le rendre plus simple. Une solution traite même en profondeur le cas (non demandé) où p et q sont multiplicativement dépendants et présente une généralisation du sujet à un certain type de séries formelles. Cette solution pourrait faire l'objet d'un article.

Dans la suite seront présentées la solution officielle, quatre solutions originales jolies, très différentes, soumises par des équipes (modifiées par moi) et finalement une solution du problème dans l'esprit de [Schae-Si]. La solution 2 procède par l'absurde en étudiant un contre-exemple tel que $\deg f - \operatorname{val} f$ est minimal. La solution 3 construit deux suites (f_n), (g_n) de polynômes satisfaisant (1) de valuation croissante et montre qu'on arrive en fait à 0. La solution 4 élimine en un premier temps tous les termes de f d'exposant divisible par p et montre qu'on arrive ainsi à $f = 0$. La solution 5 caractérise les polynômes dans l'image de l'opérateur $d(x) \mapsto d(x^p) - d(x)$ et montre que $f(x)$ est dans l'image si $f(x^q) - f(x)$ est dedans. La dernière solution montre avec de l'analyse que la solution formelle h établie dans la solution officielle est en fait un polynôme. Cette approche permet de traiter aussi le cas matriciel mentionné ci-dessus. Il serait intéressant d'en trouver une solution purement algébrique pour pouvoir généraliser à des corps de caractéristique strictement positive.

Pourquoi présenter plusieurs démonstrations ? L'énoncé est ainsi éclairé de plusieurs points de vue. Ceci illustre aussi que plusieurs méthodes d'attaquer un problème peuvent réussir si on est suffisamment persévérant.

Parmi les 14 autres équipes, plusieurs ont donné une solution partielle, parfois de manière élégante : le cas où p et q sont premiers entre eux. Ce cas est plus simple mais déjà intéressant. L'équation (1) montre alors que le degré de f est divisible par p et on peut raisonner par récurrence. Étant donné le grand nombre d'équipes ayant proposé une solution complète, et malgré la qualité de certaines de ces solutions partielles, j'ai dû à regret leur attribuer la note 0.

Solutions

Solution 1 (officielle). Si $r(x)$ est un polynôme complexe sans terme constant, alors il existe une série entière [1] $s(x)$ sans terme constant telle que $s(x^p) - s(x) = r(x)$: dans le cas où $r(x) = x^n$ avec $n \in \mathbb{N}^*$, la série $s(x) = -\sum_{k\geq 0} x^{n\,p^k}$ résout l'équation. Par linéarité, on a en général $s(x) = -\sum_{k\geq 0} r\left(x^{p^k}\right)$.

On considère à présent la série entière $h(x)$ sans terme constant telle que $h(x^p) - h(x) = f(x)$. On a alors $h(x^q) - h(x) = g(x)$. En effet, si on pose $\tilde{g}(x) = h(x^q) - h(x)$, alors le calcul donne $\tilde{g}(x^p) - \tilde{g}(x) = f(x^q) - f(x)$. La série entière $d(x) = \tilde{g}(x) - g(x)$ satisfait $d(x^p) - d(x) = 0$, donc $d(x) = d(0) = 0$.

Retenons que la série $h(x)$ est donnée par les deux formules

$$h(x) = -\sum_{k\geq 0} f\left(x^{p^k}\right) = -\sum_{k\geq 0} g\left(x^{q^k}\right). \tag{5}$$

Introduisons la notation $\mathcal{S}(r) = \{n \in \mathbb{N} \mid a_n \neq 0\}$, le "support" d'une série entière $r(x) = \sum_{n=0}^{\infty} a_n x^n$. Les formules (5) donnent

$$\mathcal{S}(h) \subset \bigcup_{k=0}^{\infty} \mathcal{S}(f)p^k \text{ et } \mathcal{S}(h) \subset \bigcup_{k=0}^{\infty} \mathcal{S}(g)q^k. \tag{6}$$

1. Il n'importe pas si la série converge ou non.

Nous montrons plus bas que l'intersection de $\bigcup_{k=0}^{\infty} \mathcal{S}(f)p^k$ et de $\bigcup_{k=0}^{\infty} \mathcal{S}(g)q^k$ est finie. Le support $\mathcal{S}(h)$ de h sera alors un ensemble fini, ce qui prouve que h est un polynôme.

Supposons par l'absurde que l'intersection a un nombre infini d'éléments. Puisque $\mathcal{S}(f)$ et $\mathcal{S}(g)$ sont finis, le principe des tiroirs montre qu'il existe des éléments $a \in \mathcal{S}(f)$, $b \in \mathcal{S}(g)$ et deux couples différents (k_1, ℓ_1), $(k_2, \ell_2) \in \mathbb{N}^2$ tels que

$$a\,p^{k_j} = b\,q^{\ell_j}, \quad j = 1, 2. \tag{7}$$

On en déduit $p^{k_2-k_1} = q^{\ell_2-\ell_1}$, ce qui contredit l'hypothèse.

Solution 2 (équipe MyCornea). On peut supposer sans perte que $p > q$. On écrit f et g sous la forme

$$\begin{array}{rcl} f(x) & = & a_C x^C + a_{C-1}x^{C-1} + \ldots + a_c x^c \\ g(x) & = & b_D x^D + b_{D-1}x^{D-1} + \ldots + b_d x^d \end{array} \tag{8}$$

avec $C, c, D, d \in \mathbb{N}$, $a_C, a_c, b_D, b_d \neq 0$ et donc $c, d > 0$. En comparant les termes de plus haut et de plus bas degré dans l'équation (1), on obtient $Cq = Dp$, $c = d$, $a_C = b_D$ et $a_c = b_d$.

On raisonne par l'absurde et on suppose que f, g satisfaisant (1) sont donnés tels qu'il n'existe pas de polynôme h satisfaisant (2) et tels que $C - c > 0$ est minimal (par suite, $D - d$ est aussi minimal). On montre d'abord que

$$pd > C \text{ et } qd > D\,, \tag{9}$$

sinon on aurait un autre contre-exemple

$$\begin{array}{rcl} \tilde{f}(x) & = & f(x) + a_c(x^{pc} - x^c) \\ \tilde{g}(x) & = & g(x) + b_d(x^{qc} - x^c) \end{array} \tag{10}$$

pour lequel cette différence serait strictement plus petite. Étudions maintenant les termes dans (1)

$$\begin{array}{llll} a_C x^{qC} + \ldots + a_k x^{qk} & +\ldots + a_d x^{qd} - a_C x^C - \ldots & -a_D x^D - .. - a_c x^c = \\ = b_D x^{pD} + \ldots + b_d x^{pd} & & -b_D x^D - \ldots - b_d x^d \end{array}\,. \tag{11}$$

Notons que qd n'est pas nécessairement plus grand que C.
À présent, soit $I = \{i \in \mathbb{N} \mid c \leq i \leq D,\ b_i \neq 0\}$. D'après l'hypothèse, I est non vide. On construit une bijection $j : I \to I$ telle que $b_i = b_{j(i)}$ pour tout $i \in I$ de la faÃ§on suivante. Pour $i \in I$ donné, il existe d'après (11) un $k \in \{c, \ldots, C\}$ tel que $pi = qk$ et $b_i = a_k \neq 0$. Si $k \leq D$, on pose $j(i) = k$. Sinon, il existe d'après (11) un $k_2 \in \{c, \ldots, C\}$ tel que $k = qk_2$ et $b_i = a_{k_2} \neq 0$. En continuant ainsi, on arrive à un $\tilde{k} \in \{c, \ldots, D\}$ et un $\alpha_i > 0$ tel que $pi = q^{\alpha_i}\tilde{k}$ et $b_i = a_{\tilde{k}}$. Comme $qc > D$, $\tilde{k}$ est unique. Alors on pose $j(i) = \tilde{k}$. Comme $\tilde{k} \in \{c, \ldots, D\}$, on a $b_i = a_{\tilde{k}} = b_{\tilde{k}} \neq 0$ et donc $\tilde{k} \in I$. On a donc bien une application $j : I \to I$. L'inégalité $qc > D$ implique aussi que j est injective : si $j(i_1) = j(i_2)$, alors i_1/i_2 doit être une puissance de q entre c/D et D/c. Comme I est de cardinal fini, j est bijective.

Notons $\beta > 0$ le cardinal de I et $\alpha = \sum_{i \in I} \alpha_i > 0$. On obtient

$$p^\beta \prod_{i \in I} i = q^\alpha \prod_{i \in I} j(i) \tag{12}$$

et donc $p^\beta = q^\alpha$, une contradiction.

Solution 3 (équipe HotMagma). On construit deux suites (f_n), (g_n) de polynômes sans terme constant satisfaisant

$$f_n(x^q) - f_n(x) = g_n(x^p) - g_n(x),\ n \in \mathbb{N} \tag{13}$$

par récurrence comme suit. D'abord $f_0 = f$ et $g_0 = g$. Ensuite, pour $n \in \mathbb{N}$, si $f_n = 0$ et donc par conséquent $g_n = 0$ (sinon le degré de $g_n(x^p) - g_n(x)$ est > 0) on pose $f_{n+1} = g_{n+1} = 0$. Si $f_n \neq 0$, alors on note $a_c x^c$, $c > 0$, le terme non nul de plus petit degré dans f_n. Par (13), le terme non nul de plus petit degré dans g_n doit être le même. Ensuite, on pose

$$f_{n+1}(x) = f_n(x) + a_c(x^{pc} - x^c),\ g_{n+1}(x) = g_n(x) + a_c(x^{qc} - x^c). \tag{14}$$

On vérifie aisément qu'on a $f_{n+1}(x^q) - f_{n+1}(x) = g_{n+1}(x^p) - g_{n+1}(x)$ et que, soit $f_{n+1} = g_{n+1} = 0$, soit les termes non nuls de plus petit degré dans f_{n+1} et g_{n+1} ont un degré plus grand que dans f_n et g_n.

On montre par récurrence qu'il existe pour tout n un polynôme h_n tel que $f(x) - f_n(x) = h_n(x^p) - h_n(x)$ et $g(x) - g_n(x) = h_n(x^q) - h_n(x)$. Si $f_n = g_n = 0$ pour un certain n, l'énoncé est donc démontré.

Supposons donc par l'absurde que $f_n \neq 0$ pour tout $n \in \mathbb{N}$. Notons alors k_n et ℓ_n le nombre de termes non nuls dans f_n et g_n respectivement. Par construction, les suites (k_n) et (ℓ_n) d'entiers positifs sont décroissantes et deviennent donc stationnaires. Soit alors $N, k, \ell \in \mathbb{N}$ tels qu'on a $k_n = k$ et $\ell_n = \ell$ pour $n \geq N$. Par hypothèse, k et ℓ sont non nuls. Pour $n \geq N$, il n'y a donc plus de compensation de termes dans le passage de n à $n+1$, c'est-à-dire que, dans le passage de f_n à f_{n+1}, le terme $a_c x^c$ de plus bas degré est *remplacé* par $a_c x^{pc}$ et simultanément le terme $b_c x^c = a_c x^c$ de plus bas degré dans g_n est remplacé par $b_c x^{qc}$.

Notons $\{a_1 x^{c_1}, \ldots, a_k x^{c_k}\}$ les termes non nuls de f_N, non nécessairement ordonnés par les degrés. Alors, comme on vient de le montrer, l'ensemble des termes de f_n, $n \geq N$, est de la forme $\{a_1 x^{c_1 p^{r_{n1}}}, \ldots, a_k x^{c_k p^{r_{nk}}}\}$, où les suites d'entiers $(r_{nj})_{n\in\mathbb{N}}$, $j = 1, .., k$ tendent vers l'infini. Si $\{b_1 x^{d_1}, \ldots, b_\ell x^{d_\ell}\}$ sont les termes non nuls de g_N, on a de même que $\{b_1 x^{d_1 p^{s_{n1}}}, \ldots, b_\ell x^{d_\ell p^{s_{n\ell}}}\}$ sont ceux de g_n, $n \geq N$ où les suites $(s_{nj})_{n\in\mathbb{N}}$, $j = 1, .., \ell$ tendent vers l'infini.

De plus, par (13) les termes de plus bas degré de f_n et g_n coïncident pour tout n. Alors, par le principe des tiroirs, on trouve deux indices $n_1 \geq N$ et $n_2 > n_1$ et $t \in \{1, \ldots, k\}$, $u \in \{1, \ldots, \ell\}$, tels que pour $i = 1, 2$ les termes de plus bas degré dans f_{n_i}, g_{n_i} sont

$$a_t x^{c_t p^{\alpha_i}} = b_u x^{d_u q^{\beta_i}} \text{ où } \alpha_i = r_{n_i t},\ \beta_i = s_{n_i u}. \tag{15}$$

Remarquons que $\alpha_2 > \alpha_1$ et $\beta_2 > \beta_1$ par construction. Les égalités (15) entraînent que

$$c_t p^{\alpha_i} = d_u q^{\beta_i},\ i = 1, 2 \tag{16}$$

et par division $p^{\alpha_2 - \alpha_1} = q^{\beta_2 - \beta_1}$, une contradiction.

Solution 4 (équipe PLouM). Dans la suite, notons $f(x) = \sum_{k\in\mathbb{N}^*} a_k x^k$, $g(x) = \sum_{k\in\mathbb{N}^*} b_k x^k$, où $a_k = 0$ et $b_k = 0$ pour tout $k \in \mathbb{N}$ sauf un nombre fini. Si f contient un terme $a_k x^k$ non nul avec k divisible par p, on peut passer à $\tilde{f}, \tilde{g}$ satisfaisant (1) tels que $\tilde{f}$ ne contient plus ce terme. Il suffit de poser $\tilde{f}(x) = f(x) - a_k(x^k - x^{k/p})$ et $\tilde{g}(x) = g(x) - a_k(x^{kq/p} - x^{k/p})$. Comme dans la solution précédente, il existe un h avec (2) si, et seulement si, il existe un $\tilde{h}$ avec les relations analogues pour $\tilde{f}$ et $\tilde{g}$: on pose $\tilde{h}(x) = h(x) + a_k x^{k/p}$.

Avec cette stratégie, on élimine d'abord tous les termes de f dont l'exposant de x est divisible par la plus grande puissance de p, etc. On peut donc supposer sans perte que f ne contient plus de termes dont l'exposant de x est divisible par p. Si $f = 0$, il n'y a rien à montrer.

Dans la suite, on suppose donc que $f \neq 0$ et que, pour tous les termes $a_k x^k$ non nuls de f, l'exposant k n'est pas divisible par p. On démontre plus bas

Lemme 1 *Pour tout $k \in \mathbb{N}^*$, si $a_k \neq 0$ ou $b_k \neq 0$, alors il existe $(\alpha, \beta) \in \mathbb{N}^2 \setminus \{(0,0)\}$ tel que $a_{kp^\alpha/q^\beta} \neq 0$ ou $b_{kp^\alpha/q^\beta} \neq 0$.*

Remarque : l'énoncé entraîne que kp^α/q^β est un entier. Néanmoins, il convient d'étendre la notion de a_r et b_r à des rationnels positifs par $a_r = b_r = 0$ si $r \notin \mathbb{N}$.

Comme $f \neq 0$, le lemme implique qu'il existe $k \in \mathbb{N}$ et une suite $((\alpha_n, \beta_n))_{n\in\mathbb{N}}$ dans $\mathbb{N}^2$, strictement croissante pour l'ordre lexicographique, tels que $a_{kp^{\alpha_n}/q^{\beta_n}} \neq 0$ ou $b_{kp^{\alpha_n}/q^{\beta_n}} \neq 0$. Or l'indépendance multiplicative implique que tous ces indices sont différents. Ceci est impossible pour les polynômes f et g.

Démonstration du lemme. Une comparaison des coefficients de x^k dans l'équation (1) implique

$$a_k - a_{k/q} = b_k - b_{k/p} \text{ pour tout } k \in \mathbb{N}^*. \tag{17}$$

Rappelons que $a_r = b_r = 0$ si $r \notin \mathbb{N}$.

Si $a_k \neq 0$, alors (17) implique $a_k - a_{k/q} = b_k$, (on a $b_{k/p} = 0$ car p ne divise pas k). Un des deux coefficients $a_{k/q}$ ou b_k doit donc être non nul. Dans le premier cas, l'énoncé est démontré.

Dans le deuxième cas, si $b_k \neq 0$, notons (17) pour pk à la place de k :

$$a_{pk} - a_{pk/q} = b_{pk} - b_k. \tag{18}$$

Au moins un des trois coefficients a_{pk}, $a_{pk/q}$ et b_{pk} doit alors être non nul ce qui démontre l'énoncé.

Solution 5 (équipe OTIS). Introduisons pour $a \in \mathbb{N}$, $a \geq 2$ les opérateurs $T_a : x\mathbb{C}[x] \to x\mathbb{C}[x]$, $T_a d(x) = d(x^a) - d(x)$. Comme démontré dans la première solution, il suffit de montrer que $f \in \operatorname{Im} T_p$. Par hypothèse, on a $T_q f \in \operatorname{Im} T_p$. Il convient donc de caractériser $\operatorname{Im} T_p$.

Lemme 2 *Soit $d \in x\mathbb{C}[x]$, $d(x) = \sum_{n\in\mathbb{N}^*} d_n x^n$. Alors $d \in \operatorname{Im} T_p$ si et seulement si $L_k d := \sum_{m=0}^{\infty} d_{kp^m} = 0$ pour tout $p \nmid k \in \mathbb{N}$.*

Remarque : ces sommes sont toutes finies puisque d est un polynôme.

Démonstration : pour $d(x) = T_p(x^r) = x^{pr} - x^r$, $r \in \mathbb{N}^*$, on a $L_k d = 0$ pour tout $p \nmid k$. Comme tout élément de $\operatorname{Im} T_p$ est combinaison linéaire de certains $T_p(x^r)$, la condition est nécessaire.

Elle est aussi suffisante. En effet soit d un polynôme tel que $L_k d = 0$ pour tout $p \nmid k \in \mathbb{N}$. Soit N tel que kp^N est plus grand que le degré de d. Alors

$$\sum_{m=0}^{\infty} d_{kp^m} x^{kp^m} = \sum_{m=0}^{N-1} d_{kp^m} (x^{kp^m} - x^{kp^N}) \in \operatorname{Im} T_p, \tag{19}$$

car $x^{kp^m} - x^{kp^N} = (x^{kp^m} - x^{kp^{m+1}}) + \ldots + (x^{kp^{N-1}} - x^{kp^N}) \in \operatorname{Im} T_p$ pour tout $m < N$. Pour conclure, il suffit d'utiliser que d est une somme finie d'expressions $\sum_{m=0}^{\infty} d_{kp^m} x^{kp^m}$, $p \nmid k$.

D'après le lemme appliqué à $d = T_q f$, $f = \sum_{n=0}^{\infty} f_n x^n$, et l'hypothèse $T_q f \in \operatorname{Im} T_p$, nous obtenons que

$$\sum_{m=0}^{\infty} f_{kp^m} = \sum_{\substack{m=0 \\ q \mid kp^m}}^{\infty} f_{kp^m/q} \quad \text{pour tout } p \nmid k. \tag{20}$$

En appliquant le lemme à nouveau, nous avons à démontrer que toutes ces sommes s'annulent.

Lemme 3 *Pour tout $p \nmid k \in \mathbb{N}$ ou bien $\sum_{m=0}^{\infty} f_{kp^m} = 0$, ou bien il existe $p \nmid \ell$ et $r \in \mathbb{N}$ tels que $q\ell = kp^r$ et $\sum_{m=0}^{\infty} f_{kp^m} = \sum_{m=0}^{\infty} f_{\ell p^m}$.*

Démonstration : si la somme n'est pas nulle, (20) implique que $\sum_{\substack{m=0 \\ q \mid kp^m}}^{\infty} f_{kp^m/q} \neq 0$ et il suffit de choisir r minimal tel que $q \mid kp^r$ et $\ell = kp^r/q$.

Ce lemme donne rapidement que toutes les sommes dans (20) sont zéro. En effet, supposons le contraire : il existe $p \nmid k$ tel que $\sum_{m=0}^{\infty} f_{kp^m} \neq 0$. Alors le lemme permet la construction d'une suite d'entiers $(k_j)_{j\in\mathbb{N}}$ strictement positifs : $k_0 = k$ et $p \nmid k_{j+1}$ tel que $qk_{j+1} = k_j p^{m_j}$ avec $m_j \in \mathbb{N}$ pour $j \in \mathbb{N}$ et $\displaystyle\sum_{m=0}^{\infty} f_{kp^m} = \sum_{m=0}^{\infty} f_{k_j p^m}$ pour tout j. Aucun des k_j ne peut être supérieur au degré de f, car alors $\sum_{m=0}^{\infty} f_{k_j p^m} = 0$, contraire à l'hypothèse. La suite $(k_j)_{j\in\mathbb{N}}$ est donc bornée. Il existe donc $j > i$ tel que $k_i = k_j$ par le principe des tiroirs. Ceci implique que $q^{j-i}k_i = p^{m_i+m_{i+1}+\ldots+m_{j-1}}k_i$, contraire à l'indépendance multiplicative.

Solution 6 ([Schae-Si]). La première étape est une variante de celle de la solution officielle : si $r(x)$ est un polynôme complexe sans terme constant, alors il existe une série entière $s(x)$ sans terme constant de rayon de convergence au moins 1 telle que $s(x^p) - s(x) = r(x)$: dans le cas où $r(x) = x^n$ avec $n \in \mathbb{N}^*$, la série $s(x) = -\sum_{k\geq 0} x^{np^k}$ résout l'équation et a le rayon de convergence 1. Par linéarité, on a en général $s(x) = -\sum_{k\geq 0} r\left(x^{p^k}\right)$.

On considère alors la série entière $h(x)$ sans terme constant de rayon de convergence ≥ 1 telle que $h(x^p) - h(x) = f(x)$. Le raisonnement de la solution officielle montre qu'elle satisfait (2).

Étape 2. Il existe une fonction entière $H(t)$ à croissance exponentielle telle que

$$H(pt) - H(t) = f(e^t) - f(1) \text{ et } H(0) = 0. \tag{21}$$

Pour montrer ceci, on développe la fonction entière $F(t) = f(e^t)$ à croissance exponentielle en une série entière $F(t) = \sum_{n=0}^{\infty} a_n t^n$. Il est connu que ses coefficients a_n satisfont une majoration $|a_n| \leq K\frac{A^n}{n!}$, $n \in \mathbb{N}$, avec certaines constantes positives A, K. L'unique série solution de l'équation (21) est $H(t) = \sum_{n=1}^{\infty} b_n t^n$ avec $b_n = \frac{a_n}{p^n-1}$. Ils satisfont aussi $|b_n| \leq K\frac{A^n}{n!}$, $n \in \mathbb{N}$, et H est donc à croissance exponentielle.

On vérifie comme avant que $H(qt) - H(t) = g(e^t) - g(1)$.

Étape 3. On montre plus bas que $h(e^t) - H(t)$ est une constante dans le demi-plan gauche $C^- = \{t \in \mathbb{C} \mid \Re\mathrm{e}\, t < 0\}$. Ceci impliquera que $h(x)$ doit être une fonction entière et la croissance exponentielle de $H(t)$ entraîne une croissance polynomiale de $h(x)$, c'est-à-dire qu'il existe des constantes $K, M > 0$ telles que $|h(x)| \leq K(1 + |x|)^M$ pour tout $x \in \mathbb{C}$. La fonction h doit donc être un polynôme.

Considérons donc la fonction $D(t) = h(e^t) - H(t)$, holomorphe dans C^-. Elle satisfait

$$D(pt) - D(t) = f(1) \text{ et } D(qt) - D(t) = g(1), \quad t \in C^-. \tag{22}$$

La fonction $E(t) = t\,D'(t)$ satisfait alors

$$E(pt) = E(t) \text{ et } E(qt) = E(t), \quad t \in C^-. \tag{23}$$

Par conséquent, la fonction $F(s) = E(e^s)$, définie dans la bande horizontale $B = \{s \in \mathbb{C} \mid \frac{\pi}{2} < \mathrm{Im}(s) < \frac{3\pi}{2}\}$, est $\log(p)$- et $\log(q)$-périodique. Comme $\log(q)/\log(p)$ est irrationnel et que F est continue, F est constante sur chaque droite horizontale, donc sur B tout entier par le principe des zéros isolés. Par définition, la fonction E aussi.

Ainsi il existe des constantes K, L telles que $D(t) = L\log t + K$, $t \in C^-$, c'est-à-dire

$$h(e^t) = H(t) + L\log t + K, \ t \in C^-. \tag{24}$$

Cette équation montre que la fonction $\tilde{H}(t) = h(e^t)$, définie sur C^-, peut être prolongée analytiquement en une fonction analytique sur $\mathbb{C} \setminus [0, \infty[$. Puisque $\tilde{H}(t)$ est une fonction $2\pi i$-périodique sur C^- à cause de la périodicité de l'exponentielle, l'unicité du prolongement analytique montre que $\tilde{H}(t)$ est une fonction entière et $t = 0$ n'est pas une singularité de $\tilde{H}(t)$. Ceci implique que $L = 0$, et que $h(e^t)$ et $H(t)$ diffèrent seulement d'une constante.

Références

[Schae-Si] Schäfke, R. and Singer, M.F., *Consistent systems of linear differential and difference equations*, Preprint, arXiv:1605.02616v1, 2016.

Reinhard Schäfke, Université de Strasbourg.

Problème 2 : Analyse

Auteure : Aline Bonami (Université d'Orléans)

Notations et but du problème

Soit $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ l'ensemble des réels modulo 2π. Soit d un entier strictement positif. Les fonctions sur $\mathbb{T}^d$ seront identifiées avec des fonctions 2π périodiques en chacune des variables. On peut écrire indifféremment $\int_{\mathbb{T}^d} f(x)\mathrm{d}x$ ou $\int_{(-\pi,+\pi)^d} f(x)\mathrm{d}x$ pour l'intégrale d'une fonction intégrable f définie sur $\mathbb{T}^d$. Pour $p \geq 1$, l'espace $L^p(\mathbb{T}^d)$ est l'espace des (classes de) fonctions f mesurables 2π périodiques en chacune des variables telles que

$$\|f\|_p^p := \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} |f(x)|^p \,\mathrm{d}x < \infty.$$

La suite des coefficients de Fourier d'une fonction $f \in L^1(\mathbb{T}^d)$ est définie par

$$\hat{f}(n) := \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} f(x) e^{-in\cdot x} \,\mathrm{d}x, \qquad n \in \mathbb{Z}^d.$$

Ici $n \cdot x$ est le produit scalaire $\sum n_j x_j$. On notera $|x|$ la norme euclidienne d'un vecteur x de $\mathbb{R}^d$. La définition de coefficient de Fourier s'étend aux mesures de Borel bornées sur $\mathbb{T}^d$. Pour une telle mesure μ, on pose

$$\hat{\mu}(n) := \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} e^{-in\cdot x} \,\mu(\mathrm{d}x), \qquad n \in \mathbb{Z}^d.$$

Soit B un espace de Banach de fonctions intégrables sur $\mathbb{T}^d$ dans lequel les polynômes trigonométriques sont denses. On dira que la suite $(m(n))_{n\in\mathbb{Z}^d}$ est un **multiplicateur de Fourier** de B, ou plus simplement, un multiplicateur, s'il existe une constante C telle que, pour tout polynôme trigonométrique f, le polynôme trigonométrique

$$T_m f(x) := \sum_{n\in\mathbb{Z}^d} m(n)\hat{f}(n)e^{in\cdot x}$$

satisfait l'inégalité

$$\|T_m f\|_B \leq C\,\|f\|_B.$$

La norme du multiplicateur, notée $\|m\|$, est la borne inférieure des constantes C pour lesquelles une telle inégalité a lieu.

Lorsque B est l'espace $L^1(\mathbb{T}^d)$, l'espace des multiplicateurs coïncide avec l'espace des coefficients de Fourier des mesures bornées. La mesure μ correspondant au multiplicateur m est telle que $T_m f = \mu \star f$ (produit de convolution).

L'espace de Sobolev $W^{1,1}(\mathbb{T}^d)$ est le complété de l'espace des polynômes trigonométriques pour la norme

$$\|f\|_{W^{1,1}(\mathbb{T}^d)} := \|f\|_1 + \|\nabla f\|_1.$$

On se propose de montrer qu'alors que l'espace des multiplicateurs de Fourier de $W^{1,1}(\mathbb{T})$ coïncide avec l'espace des multiplicateurs de $L^1(\mathbb{T})$, ce n'est plus le cas en dimension supérieure. La recherche de contre-exemple sera basée sur l'inégalité de Gagliardo-Nirenberg.

A. Résultats préliminaires

1. Montrer que la suite des coefficients de Fourier d'une mesure bornée est un multiplicateur de $W^{1,1}(\mathbb{T}^d)$ quelle que soit la dimension d.
2. Montrer que si m est un multiplicateur de $W^{1,1}(\mathbb{T}^d)$,

$$\sup_{n\in\mathbb{Z}^d} |m(n)| \leq \|m\|.$$

3. Montrer qu'il n'y a pas d'autres multiplicateurs de $W^{1,1}(\mathbb{T}^d)$ que les suites de coefficients de Fourier des mesures bornées lorsque la dimension est 1.
4. Montrer qu'il suffit de trouver un contre-exemple en dimension 2 pour construire un contre-exemple en toute dimension $d > 1$.

B. Le contre-exemple

1. On pourra montrer ou admettre l'inégalité de Gagliardo-Nirenberg pour $W^{1,1}(\mathbb{T}^2)$: il existe une constante C telle que, pour tout polynôme trigonométrique,

$$\|f\|_2 \leq C\,\|f\|_{W^{1,1}(\mathbb{T}^2)}.$$

2. On pourra admettre l'inégalité de Khintchine : soit X_j une suite de variables aléatoires indépendantes de Bernoulli, c'est-à-dire prenant les valeurs 1 et -1 avec probabilité 1/2. Il existe une constante C telle que pour toute suite finie a_n on ait

$$(\sum |a_j|^2)^{1/2} \leq C\,\mathbb{E}(|\sum a_j X_j|).$$

3. Montrer l'existence d'une constante C telle que, quelle que soit la suite $(\varepsilon_n)_{n\in\mathbb{Z}^d}$ de coefficients égaux à ± 1, la suite $\frac{\varepsilon_n}{(|n|^2+1)^{1/2}}$ est un multiplicateur de norme bornée par C.
4. En déduire qu'il existe des multiplicateurs qui ne sont pas les coefficients de Fourier d'une mesure bornée.

Rapport

Le sujet se présentait comme un problème usuel, avec quelques questions élémentaires dans une première partie, puis une accélération dans la seconde dans laquelle on regroupait les outils nécessaires pour arriver à une question finale dans laquelle on n'était plus guidé.

La question qui y est abordée est celle des opérateurs de convolution qui préservent l'espace de Sobolev $W^{1,1}$ des fonctions intégrables dont les dérivées partielles sont intégrables. Il est plus simple de le faire dans le cadre périodique, ce qui avait l'avantage supplémentaire qu'il n'y avait pas directement de littérature disponible. Il s'agissait de montrer qu'alors qu'en dimension 1 les seules convolutions qui préservent l'espace sont les convolutions avec des mesures bornées, il n'en est rien en dimension supérieure. Pour obtenir un contre-exemple, on montrait qu'il était

suffisant de considérer la dimension 2 pour laquelle l'espace de Sobolev se plonge dans L^2, ce qui permettait d'avoir toute une famille de telles convolutions. Puis on utilisait les inégalités de Khintchine pour en déduire l'existence du contre-exemple. C'est dans cette dernière partie que les étudiants se trouvaient livrés à eux-même, en particulier pour réaliser qu'il fallait d'une façon ou une autre raisonner par l'absurde pour dire que s'il n'y avait pas de contre-exemple on aurait une borne uniforme pour les normes des mesures. Seule la meilleure copie, qui était excellente, a fait appel aux théorèmes d'analyse fonctionnelle à cet endroit. Les copies qui ont été retenues, au nombre de 6, sont celles qui ont compris l'ensemble de la démarche, les deux meilleures notes étant réservées à celles qui se sont réellement attaquées à cette dernière question.

Le fait qu'il y ait de tels contre-exemples a été trouvé par Poornima dans les années 1980 en utilisant les *non L^1 inequalities* de D. Ornstein, c'est-à-dire des exemples de distributions d'ordre 1 dont toutes les dérivées partielles sont d'ordre 1 (ceci évidemment lorsque la dimension est supérieure à 1). J'ai peu après montré avec elle, toujours en utilisant le théorème d'Ornstein, qu'aucun opérateur d'intégrale singulière ne préserve l'espace $W^{1,1}(\mathbb{R}^d)$. J'aurais aimé, et j'aimerais toujours, montrer que les convoluteurs des espaces $W^{k,1}$ diffèrent, non pas uniquement entre $k = 0$ et $k > 0$, mais pour toutes les valeurs de k. L'espace $W^{1,1}(\mathbb{R}^d)$ est un espace dont on sait bien qu'il réserve des mystères, tout comme l'espace BV. Curieusement, la construction d'Ornstein a attiré l'attention de plusieurs auteurs ces dernières années. Je suis revenue sur le sujet ces dernières années grâce à la lecture d'un article récent de deux collègues polonais, Kazaniecki et Wojciechowski, dans laquelle ils vont plus loin que nous n'étions allées dans les années quatre-vingts et montrent que les transformées de Fourier des convoluteurs de $W^{1,1}(\mathbb{R}^d)$ sont continues. Le texte du problème est issu de réflexions récentes. La construction qui est donnée dans le problème permet de montrer qu'il y a d'autres opérateurs de convolution que ceux qui sont donnés par des distributions d'ordre 1 dont les dérivées partielles sont aussi d'ordre 1, autrement dit que Poornima n'avait pas trouvé tous les opérateurs de convolution de $W^{1,1}(\mathbb{R}^d)$.

Certaines copies, par ailleurs soignées, ont montré un réel manque de culture sur les séries de Fourier. La démarche de la dernière question est évidemment naturelle si on sait comment démontrer qu'il n'y a pas convergence des séries de Fourier dans L^1. Plus élémentaires, les manipulations sur les séries doubles n'étaient pas forcément aisées.

Solution

Donnons la démonstration d'un résultat donné dans l'énoncé : *Lorsque B est l'espace $L^1(\mathbb{T}^d)$, l'espace des multiplicateurs coïncide avec l'espace des coefficients de Fourier des mesures bornées. La mesure μ correspondant au multiplicateur m est telle que $T_m f = \mu * f$.*

En effet, le fait que la transformée de Fourier d'une mesure soit un multiplicateur est immédiat : la convolution se transforme en multiplication. Réciproquement, on teste le multiplicateur sur une suite de polynômes de norme uniformément bornée qui forment une identité approchée, par exemple sur le produit tensoriel de noyaux de Fejer. On trouve que $\|T_m f_N\|_1 \leq C$. Il existe donc une mesure μ qui est la limite faible* de $T_m f_N$. On prend maintenant pour fonctions tests les exponentielles $e^{im.x}$ pour voir que $m(n)$ est la suite des coefficients de Fourier de μ.

A. Résultats préliminaires

1. Pour montrer que la suite des coefficients de Fourier d'une mesure bornée est un multiplicateur de $W^{1,1}(\mathbb{T}^d)$, on utilise les relations de commutation entre dérivation et convolution.

2. Pour démontrer l'inégalité $\sup_{n\in\mathbb{Z}^d} |m(n)| \leq \|m\|$, on teste le multiplicateur sur l'exponentielle $e^{in\cdot x}$.

3. Cas de la dimension 1. Soit $d = 1$ et m un multiplicateur de $W^{1,1}(\mathbb{T})$. Nous allons montrer que m est aussi un multiplicateur de $L^1(\mathbb{R})$. Soit tout d'abord f un polynôme trigonométrique de moyenne nulle. La fonction $g(x) := \int_0^x f(t)dt$ est 2π périodique. C'est à nouveau un polynôme trigonométrique, de dérivée f, qui est borné par $2\pi\|f\|_1$. La dérivée de $T_m g$ s'écrit
$$(T_m g)'(x) = \sum_{n\neq 0} m(n)\widehat{f}(n)e^{inx}.$$
On reconnaît la fonction $T_m f$. Puisque m est un multiplicateur de $W^{1,1}(\mathbb{T})$, on a donc l'inégalité
$$\|T_m f\|_1 \leq \|m\|(\|g\|_1 + \|f\|_1) \leq (1+2\pi)\|m\|\|f\|_1.$$
Soit maintenant h un polynôme trigonométrique quelconque, qu'on peut écrire sous la forme $h = \widehat{h}(0) + f$, avec f de moyenne nulle. Alors
$$\|T_m h\|_1 \leq |m(0)|\|\widehat{h}(0)| + (1+2\pi)\|m\|\|f\|_1 \leq (3+4\pi)\|m\|\|h\|_1.$$
On a utilisé le fait que $|\widehat{h}(0)| \leq \|h\|_1$, la question précédente pour majorer $|m(0)|$, et le fait que $\|f\|_1 \leq 2\|h\|_1$. Le multiplicateur m, qui est aussi un multiplicateur de $L^1(\mathbb{T})$, est donné par la suite des coefficients de Fourier d'une mesure bornée.

4. On suppose qu'on a un multiplicateur m de $W^{1,1}(\mathbb{T}^2)$ et on va construire un multiplicateur de $W^{1,1}(\mathbb{T}^2 \times \mathbb{T}^d)$. On pose
$$\widetilde{m}(n,n') := \begin{cases} 0 & \text{si } n' \neq 0 \\ m(n) & \text{si } n' = 0 \end{cases}.$$
Ici n est dans $\mathbb{Z}^2$ et n' dans $\mathbb{Z}^d$. On vérifie aisément que, si f est un polynôme trigonométrique de $\mathbb{T}^d$,
$$T_{\widetilde{m}} f(x,x') = T_m g(x),$$
où g est le polynôme trigonométrique sur $\mathbb{T}^2$ défini par
$$g(x) := \sum_{n\in\mathbb{Z}^2, n'=0} \widehat{f}(n,n')e^{i(n\cdot x + n'\cdot x')}.$$
g est donné par la multiplication des coefficients de Fourier de f par ceux de la mesure $\delta_0(x)\otimes 1(x')$, donc par la convolution avec une mesure de norme 1. On en déduit que
$$\|g\|_{W^{1,1}(\mathbb{T}^2)} \leq \|f\|_{W^{1,1}(\mathbb{T}^2\times\mathbb{T}^d)},$$
et que $\widetilde{m}$ est un multiplicateur de $W^{1,1}(\mathbb{T}^2\times\mathbb{T}^d)$. Supposons que $\sum \widetilde{m}(n)e^{in\cdot x}$ soit la série de Fourier d'une mesure bornée de $\mathbb{T}^2\times\mathbb{T}^d$. Celle-ci ne dépend pas de la variable dans $\mathbb{T}^d$. C'est donc une mesure bornée sur $\mathbb{T}^2$, ce qui mène à une contradiction si le multiplicateur avec lequel on travaille est un contre-exemple.

B. Le contre-exemple

1. La démonstration classique, valable pour les variétés, commence par une partition de l'unité. Voici ici une démonstration peut-être techniquement plus compliquée, mais sans cette étape. Considérons tout d'abord le cas de $\mathbb{T}$ et montrons qu'il existe une constante C telle que, si f est un polynôme trigonométrique,
$$\|f\|_\infty \leq C\|f\|_{W^{1,1}(\mathbb{T})}. \tag{25}$$
Il suffit d'écrire $f(x) = \int_0^x f'(t)\mathrm{d}t + A$. La constante A est fixée par le fait que
$$\int_0^{2\pi}\int_0^x f'(t)\,\mathrm{d}t\mathrm{d}x + 2\pi A = \int_0^{2\pi} f(x)\,\mathrm{d}x.$$

On en déduit l'inégalité avec $C = 4\pi$. Soit maintenant f un polynôme trigonométrique de deux variables. On utilise l'inégalité précédente séparément en chacune des deux variables. Il en résulte qu'on a $|f(x_1,x_2)|^2 \leq C^2\psi(x_1)\phi(x_2)$, où

$$\phi(x_2) = \int_{\mathbb{T}} |f(x_1,x_2)|\,\mathrm{d}x_1 + \int_{\mathbb{T}} |\partial_{x_1} f(x_1,x_2)|\,\mathrm{d}x_1$$

et ψ est défini de façon analogue. Donc

$$\int_{\mathbb{T}^2} |f(x_1,x_2)|^2\,\mathrm{d}x_1\mathrm{d}x_2 \leq C^2 \int_{\mathbb{T}} |\psi(x_1)|\,\mathrm{d}x_1 \times \int_{\mathbb{T}} |\phi(x_2)|\,\mathrm{d}x_2.$$

On en déduit que

$$\int_{\mathbb{T}^2} |f(x_1,x_2)|^2\,\mathrm{d}x_1\mathrm{d}x_2 \leq C^2 \left(\int_{\mathbb{T}^2} (|f(x)| + |\nabla f|)\,\mathrm{d}x \right)^2,$$

ce qu'on voulait démontrer.

2. La démonstration est partout : on commence par démontrer que la norme L^4 est majorée, à une constante près, par la norme L^2, puis on utilise l'inégalité de Hölder.

3. L'inégalité de Gagliardo-Nirenberg suivie de l'identité de Plancherel nous permettent d'affirmer l'existence d'une constante C telle que, pour tout polynôme trigonométrique f,

$$\left(\sum_{n\in\mathbb{Z}^2} |\widehat{f}(n)|^2 \right)^{1/2} \leq C\|f\|_{W^{1,1}(\mathbb{T}^2)}.$$

Soit $m(n) = \frac{\varepsilon_n}{(|n|^2+1)^{1/2}}$. Il en résulte que

$$\left(\sum_{n\in\mathbb{Z}^2} (|n_1|^2 + |n_2|^2 + 1)|m(n)\widehat{f}(n)|^2 \right)^{1/2} \leq C\|f\|^2_{W^{1,1}(\mathbb{T}^2)}.$$

Comme $\widehat{\partial_{x_j} f} = in_j\widehat{f}$, on reconnait dans le membre de gauche $\|T_m f\|_2^2 + \|\partial_{x_1} T_m f\|_2^2 + \|\partial_{x_2} T_m f\|_2^2$. En utilisant les inclusions entre espaces L^p on en déduit que

$$\|T_m f\|_{W^{1,1}(\mathbb{T}^2)} \leq C\|f\|_{W^{1,1}(\mathbb{T}^2)},$$

ce qu'on voulait démontrer.

4. On montre par l'absurde qu'il existe des multiplicateurs qui ne sont pas les coefficients de Fourier d'une mesure bornée. Supposons que tout multiplicateur m soit la suite des coefficients de Fourier d'une mesure $S(m)$. Il est facile de voir que les deux espaces, espace des multiplicateurs et espace des mesures bornées, sont des espaces de Banach. De plus l'opérateur S est de graphe fermé : si m_k tend vers m et $S(m_k)$ tend vers μ, alors $\widehat{\mu}(n) = \lim m_k(n) = m(n)$. On a utilisé le fait que pour tout n l'application qui à un multiplicateur fait correspondre sa valeur en n est continue (section 2, question 2). Il existe donc une constante C telle que, pour tout polynôme trigonométrique f de norme 1 dans $L^1(\mathbb{T}^2)$ et pour toute suite ε prenant les valeurs ± 1,

$$\int_{\mathbb{T}^2} \left| \sum_{n\in\mathbb{Z}^2} \frac{\varepsilon_n}{(|n|^2+1)^{1/2}} \widehat{f}(n) e^{in\cdot x} \right| \mathrm{d}x \leq C.$$

On prend pour suite ε_n la suite $X_n(\omega)$, où X_n est une suite de v.a. de Bernoulli indépendantes sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$. On intègre par rapport à ω et on échange les ordres d'intégration. L'inégalité de Khintchine nous permet de dire qu'il existe une constante C telle que, pour tout polynôme trigonométrique f de norme 1 dans $L^1(\mathbb{T}^2)$

$$\sum_{n\in\mathbb{Z}^2} \frac{|\widehat{f}(n)|^2}{|n|^2+1} \leq C.$$

Il suffit par exemple de prendre pour f des produits de noyaux de Cesàro en chacune des variables pour obtenir une contradiction avec le fait que

$$\sum_{n\in\mathbb{Z}^2} \frac{1}{|n|^2+1} = \infty.$$

Aline Bonami, Université d'Orléans

Problème 3 : Combinatoire-cryptographie

Auteur : Olivier Poisson, (Aix-Marseille Université)

On se donne des entiers $p \geq 1$, $n \geq 2$, $k_i \geq 1$ pour $i = 1, \ldots, n$, tels que

$$k_1 + \ldots + k_n = 2p\,.$$

Soit l'ensemble $\mathcal{S}$ des solutions de l'équation

$$x_1 + \ldots + x_n = p\,,\ 0 \leq x_i \leq k_i\,.$$

Il s'agit de donner un minorant de $c_p := \#\mathcal{S}$ (cardinal de $\mathcal{S}$), notamment dans le cas p grand, qui soit explicite selon $k_1, \ldots, k_n$. Montrer qu'on a par exemple

$$c_p \geq C_n \frac{\Pi_{i=1}^n k_i}{\sqrt{\sum_{i=1}^n k_i^2}}, \tag{26}$$

pour une constante $C_n > 0$, dont on pourra donner une expression ou un moyen de calcul.

Il n'est pas interdit de proposer une meilleure minoration de c_p, ou même un équivalent simple quand p est grand.

Rapport

(Ce problème m'a été proposé il y a une dizaine d'années et à peu près tel quel par un ingénieur d'Intertechnique (actuellement Zodiac ou Safran), sans la relation entre c_p et C_n. Celui-ci a malheureusement perdu la trace de son origine.)

Dix-neuf équipes ont rendu une copie. La plupart étaient acceptables. J'ai dû mettre malheureusement o à quelques copies parce qu'elles ne donnaient de minorant c_p que pour $n \leq 3$ ou bien parce qu'il était moins bon que la forme demandée (C_n dépendant de p) Cette forme peut d'ailleurs être généralisée sans dommage en $c_p = C_{N,n} \frac{\Pi_{i=1}^n k_i}{N(k)}$, avec $k = (k_1, \ldots, k_n)$ et N une norme

sur $\mathbb{R}^n$. En particulier, $N(k) = \|k\|_\infty := \sup_i |k_i|$ est apparue de nombreuses fois, avec $\Pi_{i=1}^n k_i$ remplacé par $\Pi_{i=1}^n (k_i + 1)$ (ce qui n'est pas pire) ; d'où un majorant sous la forme $c_p = C_{\|\|_\infty, n} \Pi_{i=1}^{n-1} k_i$, lorsque $k_1 \leq k_2 \ldots \leq k_n$. Ce résultat est rapidement obtenu lorsque l'on s'aperçoit de la croissance de la suite $m \mapsto \#\mathcal{S}(m)$, $0 \leq m \leq p$, avec $\mathcal{S}(m)$ l'ensemble de solutions entières de l'équation

$$x_1 + \cdots + x_n = m \quad \text{et } 0 \leq x_i \leq k_i, \quad \forall i.$$

La preuve de la croissance de $\#\mathcal{S}(\cdot)$ est faite par récurrence, mais il vaut mieux faire cette récurrence sur n que sur p.

Pour déterminer une valeur de C_n, deux approches ont été utilisées principalement.
D'abord, la méthode géométrique, qui compare $\mathcal{S}(p) = \mathcal{S}$ avec l'intersection de l'hyperplan affine $\{x \in \mathbb{R}^n ; < x, u >= p\}$, $u = (1, \ldots, 1)$ et du pavé $\Pi_{i=1}^n [0, k_i]$. Cette méthode permet de donner une estimation asymptotique de c_p par exemple si $k_i/p \to q_i \in]0, 2[$ lorsque $p \to \infty$. En effet, on peut comparer c_p avec $C'_n p^{n-1}$ où C'_n dépend de mesures (de Hausdorf) d'intersection d'hyperplans avec l'hypercube unité C^n dans $\mathbb{R}^n$ (les hyperplans coupant l'hypercube par le centre). Il est possible de donner des estimations simples de la mesure de ce genre d'hypersurface en comparant par exemple l'hypercube C^n avec des boules euclidiennes centrées en o. Une autre possibilité, moins élémentaire, est fournie dans une copie qui utilise le fait que le $(n-1)$-volume de l'intersection d'un hyperplan orthogonal à u avec un pavé de $\mathbb{R}^n$ (coupé en son centre) est minimale lorsque l'hyperplan est orthogonal à e_i (i^{eme}) vecteur de la base canonique de $\mathbb{R}^n$ tel que $k_i = \max_j k_j$. (La copie se réfère à un résultat dans le livre d'Alexander Koldobsky : "Fourier Analysis in Convex Geometry", AMS, 2014. J'aurai préféré plus de détails, mais le problème, bien que facilement posé, semble compliqué à résoudre).
Ensuite, la deuxième méthode, probabiliste, part de la considération de n v.a.r X_i, i.i.d, de loi uniforme sur $[0, k_i]$, $i = 1, \ldots, n$. Cette approche me parait trop sophistiquée lorsqu'elle se limite à démontrer la croissance de $\#\mathcal{S}(\cdot)$. Sinon, à partir de ce point, et en utilisant par exemple l'inégalité de Bien Aymé-Tchebytchev (ou même de Hoeffding, sans doute plus précise), elle fournit une très bonne estimation de $C_{\|\cdot\|_2, n}$, indépendante de n (sinon, à partir de $C_{\|\cdot\|_\infty, n}$ et de l'inégalité de Cauchy-Schwarz, on obtient $C/\sqrt{n}$, C : numérique), avec $\|\cdot\|_2$ la norme euclidienne sur $\mathbb{R}^n$. A noter que l'utilisation de Hoeffding seule, sans la propriété de croissance de $\#\mathcal{S}(\cdot)$, donne un résultat insuffisant sur C_n.

Un bon lot de copies s'est attaqué à la détermination d'un équivalent de c_p lorsque p est grand, soit, comme déjà commenté, en prenant la méthode géométrique, qui me parait la plus élégante, soit par le théorème du crible. Malgré quelques erreurs mineures ou des preuves un peu rapides, il est assez satisfaisant de voir apparaitre un résultat substantiel.

Pour finir, j'ai été agréablement surpris de recevoir de nombreuses copies suffisamment riches, certaines m'ayant permis d'avoir un nouveau regard sur le problème que j'ai soumis.

Solution

Cas $n = 2$

Ce cas est trivial. Puisque $k_1 + k_2 = 2p$, on peut supposer $0 \leq k_1 \leq p \leq k_2 \leq 2p$. Donc si $x_1 + x_2 = p$, $x_i \in [0, k_i] \cap \mathbb{N}$, alors en se fixant $x_1 \in [0, k_1] \cap \mathbb{N}$, on obtient $x_2 = p - x_1 \in [0, p] \subset [0, k_2]$. Il y a $1 + k_1 = 1 + \min(k_1, k_2)$ solutions de la forme $(x_1, p - x_1)$, $0 \leq x_1 \leq k_1$. Donc $c_p = 1 + \min(k_1, k_2)$.

Méthode Géométrique

On suppose $n \geq 3$, $p \geq 2$. On peut toujours supposer $k_1 \leq k_2 \leq \cdots \leq k_n$. Rappelons qu'on a supposé $k_i > 0$ pour tout i. En effet, si $k_1 = \cdots = k_r = 0$ et $k_{r+1} > 0$, alors on se ramène au même problème avec $n - r$ à la place de n en ne considérant que les inconnues $x_{r+1}, \ldots, x_n$.

Préliminaires. On considère le pavé $P_n = [0,k_1]\times\cdots[0,k_n]$ et $K_n = P_n \cap \mathbb{N}^n$. Le symbole $\preceq$ désigne la relation d'ordre (partiel) suivant sur $\mathbb{N}^n$:

$$x \preceq y \Leftrightarrow x_i \leq y_i \quad \forall i.$$

Donc $K_n = [[0_{\mathbb{R}^n}, k]]$ comme intervalle pour cet ordre $\preceq$, où $k = (k_1,\ldots,k_n)$. Les entiers $x_1 \ldots x_n$ forment une solution de

$$\sum_{i=1}^{n} x_i = p\,,\ 0 \leq x_i \leq k_i,$$

si et seulement si le vecteur $x = (x_1,\ldots,x_n) \in K_n$ satisfait $< x,u >= p$ où $u = (1,\ldots,1) \in \mathbb{R}^n$ et $<,>$ désigne le produit scalaire euclidien de $\mathbb{R}^n$. [Remarquons alors qu'il existe au moins un point dans S. En effet, l'application $(K_n,\preceq) \ni x \mapsto < x,u >$ est croissante de 1 selon chaque coordonnée x_i, avec $< 0_{\mathbb{R}^n}, u >= 0$, $< k,u >= 2p$ (et les points $0_{\mathbb{R}^n}$, k forment les extrêmités de l'intervalle $K_n = [[0_{\mathbb{R}^n}, k]]$), donc on peut utiliser un théorème de valeurs intermédiaires discret.] L'ensemble des solutions cherché est donc $\mathcal{S} = K_n \cap E_u^p$, où nous notons généralement par E_v^a l'hyperplan affine dans $\mathbb{R}^n$ de direction $v^\perp$ tel que $x \in E_v^a \Longleftrightarrow < x,v >= a$, pour $v \in \mathbb{R}^n \setminus \{0\}$ et $a \in \mathbb{R}$. Notamment, E_u^p passe par $\frac{1}{2}k$, point qui peut ne pas être dans K_n, mais appartient à P_n : c'est son centre. L'hyperplan $E_u^p - k/2$ de direction $u^\perp$ passe par 0, puisque $< k/2, u >= \sum_{i=1}^n k_i/2 = p$ par hypothèse. Nous avons donc $E_u^0 = E_u^p - k/2$.

Notons par $H(x,r) = x + [0,r]^n \subset \mathbb{R}^n$ l'hypercube basé en $x \in \mathbb{R}^n$ et de longueur de coté $r > 0$. Notons aussi par $H_0 = H(0,1)$ l'hypercube usuel $[0,1]^n$ centré en $u/2$, puis $H_c = H_0 - u/2 = H(-u/2,1)$ l'hypercube centré en 0, de longueur de coté 1. Nous notons par vol_j le volume d'une surface de dimension j. Nous posons

$$A_n(v) = \mathrm{vol}_{n-1}(H_c \cap E_v^0), \quad v \in \mathbb{R}^n \setminus \{0\}.$$

Ce nombre ne dépend que de n et de v. Nous l'estimerons par la suite. Il est possible aussi de le garder tel quel ou d'en donner des approximations numériques si n n'est pas très grand.

Minoration de $c_p = \#\mathcal{S}$.

A chaque solution $x \in \mathcal{S}$ on fait correspondre l'hypercube $K^x = x + 2H_c$ de centre x, de coté de longueur 2. Ces hypercubes K^x constituent une famille finie F de cardinal $\#F = c_p$, se chevauchent [2], et remplissent le volume $S_n = \cup_{x\in\mathcal{S}} K^x = \cup_{h\in F} h$. Ce volume S_n contient $S' := P_n \cap E_u^p$. En effet, soit $y = (y_1,\ldots,y_n) \in S'$. Posons $y^\pm = (y_1^\pm,\ldots,y_n^\pm)$ avec y_i^- = plus grand entier $\leq y_i$ et y_i^+ = plus petit entier $\geq y_i$. Donc $y^\pm \in K - n$ et $< y^-, u >\preceq< y, u >= p \preceq< y^+, u >$. Par théorème des valeurs intermédiaires, il existe donc $x \in [[y^-, y^+]] \subset K_n$ tel que $< x,u >= p$. De plus, pour tout i on a

$$|y_i - x_i| = \max(y_i - x_i, x_i - y_i) \leq \max(y_i^+ - x_i, x_i - y_i^-) \leq y_i^+ - y_i^- \leq 1.$$

Donc $y \in K^x$, ce qui prouve que $S' \subset S_n$.
Par conséquent, nous avons

$$S' \subset S_n \cap E_u^p = (\cup_{h\in F} h) \cap E_u^p = \cup_{h\in F}(h \cap E_u^p),$$

puis

$$\mathrm{vol}_{n-1}(S') \leq \sum_{h\in F} a_u(h),$$

où on a posé $a_u(h) := \mathrm{vol}_{n-1}(h \cap E_u^p)$. Remarquons que, pour $x \in \mathcal{S}$, nous avons

$$\begin{aligned} E_u^0 - x + k/2 &= \{y \in \mathbb{R}^n; y + x - k/2 \in E_u^0\} = \{y \in \mathbb{R}^n; < y + x - k/2, u >= 0\} = \{y \in \mathbb{R}^n; < y, u > \\ &= < -x + k/2, u >= 0\} = E_u^0. \end{aligned}$$

2. malheureusement ! il y a donc perte d'optimalité

Donc,

$$K^x \cap E_u^p = (x + 2H_c) \cap (E_u^0 + k/2) = (2H_c \cap (E_u^0 - x + k/2)) + x = (2H_c \cap E_u^0) + x,$$

puis, par propriété d'invariance du volume par translation,

$$a_u(K^x) = \mathrm{vol}_{n-1}(K^x \cap E_u^p) = \mathrm{vol}_{n-1}(2H_c \cap E_u^0), \quad x \in \mathcal{S}.$$

Ceci montre que $a_u(h) \equiv a_n$ ne dépend pas de $h \in F$. De plus, par homothétie de rapport 2, nous avons

$$a_n = \mathrm{vol}_{n-1}(2H_c \cap E_u^0) = \mathrm{vol}_{n-1}(2(H_c \cap E_u^0)) = 2^{n-1}\mathrm{vol}_{n-1}((H_c \cap E_u^0)) \equiv 2^{n-1}A_n(u).$$

Ainsi nous avons

$$\mathrm{vol}_{n-1}(S') \leq \#F\, a_n = c_p\, 2^{n-1} A_n(u).$$

Déterminons $\mathrm{vol}_{n-1}(S')$, $S' = P_n \cap E_u^p$. Considérons le changement de variable $y = f(x)$ avec $y_i = (x_i - k_i/2)/k_i$, $1 \leq i \leq n$. Nous avons $f(k/2) = 0$, $f(P_n) = H_c$ et $f(E_u^p) = E_k^0$, puis $f(S') = H_c \cap E_k^0$, et donc, par un calcul usuel,

$$\mathrm{vol}_{n-1}(S') = |Jac(f)|\frac{\|u\|_2}{\|k\|_2}\mathrm{vol}_{n-1}(f(S')) = \frac{\Pi_{i=1}^n k_i}{\sqrt{\sum_{i=1}^n k_i^2}}\sqrt{n}A_n(k).$$

D'où :

$$c_p \geq \frac{\mathrm{vol}_{n-1}(S')}{2^{n-1}A_n(u)} = \frac{\Pi_{i=1}^n k_i}{\sqrt{\sum_{i=1}^n k_i^2}}\frac{\sqrt{n}}{2^{n-1}}\frac{A_n(k)}{A_n(u)}.$$

Pour une majoration de $A_n(u)$, on peut par exemple remarquer que H_c est inclus dans la boule euclidienne fermée de $\mathbb{R}^n$, notée $B'_n(0, r_n)$, centrée en o et de rayon $\frac{1}{2}\|u\|_2 = \sqrt{n}/2$. Comme l'intersection de $B'_n(0, \sqrt{n}/2)$ avec E_u^0 est géométriquement la boule euclidienne $B'_{n-1}(0, \sqrt{n}/2)$ dans $\mathbb{R}^{n-1}$ nous avons donc

$$A_n(u) \leq \mathrm{vol}_{n-1}(B'_n(0, \sqrt{n}/2) \cap E_u^0) = \mathrm{vol}_{n-1}(B'_{n-1}(0, \sqrt{n}/2)) = (\sqrt{n}/2)^{n-1}\, V_{n-1}\,, \tag{27}$$

où V_{n-1} désigne le volume de la boule unité euclidienne dans $\mathbb{R}^{n-1}$:

$$V_{2j} = \frac{\pi^j}{j!}\,,\ V_{2j+1} = \frac{\pi^j}{(j+\frac{1}{2})(j-\frac{1}{2})\cdots\frac{1}{2}}.$$

Pour une minoration (grossière) de $A_n(k)$, nous voyons que H_c contient la boule euclidienne unité $B'_n(0, 1/2)$. Comme l'intersection de $B'_n(0, 1/2)$ avec E_k^0 est géométriquement la boule euclidienne $B'_{n-1}(0, 1/2)$ dans $\mathbb{R}^{n-1}$, nous avons donc

$$A_n(k) \geq \mathrm{vol}_{n-1}(B'_{n-1}(0, 1/2)) = 2^{1-n}V_{n-1}. \tag{28}$$

En conséquence nous obtenons

$$c_p \geq C_n \frac{\Pi_{i=1}^n k_i}{\sqrt{\sum_{i=1}^n k_i^2}}, \quad C_n = \frac{\sqrt{n}}{4^{n-1}}\frac{V_{n-1}}{A_n(u)} \geq \frac{1}{2^{n-1}(\sqrt{n})^{n-2}}.$$

Olivier Poisson, Aix-Marseille Université.

Problème 4 : Géométrie

Auteur : Nicolas Juillet (Université de Strasbourg)

Sur un plan pavé par des triangles équilatéraux on déplace un solide régulier (à faces triangulaires) dont les faces ont la même taille que les dalles du pavage. Le solide est posé sur une seule dalle et pour se déplacer sur les dalles voisines il bascule sur l'arête commune aux dalles d'avant et d'après le mouvement.

À partir d'une position initiale, quelles sont les positions (dalle et orientation du solide) atteignables en se déplaçant ? On étudiera les solides réguliers suivants :

1. le tétraèdre,
2. l'octaèdre,
3. l'icosaèdre.
4. On se pose la question supplémentaires des positions atteignables lorsqu'un tétraèdre roule sur un icosaèdre.

Rapport

Le principe était de proposer un problème inspiré de la géométrie sous-riemannienne qui puisse avoir une formulation suffisamment simple pour se prêter à l'expérimentation. On peut voir ce problème comme un pendant discret d'un exemple classique, celui d'une sphère roulant sans glisser sur un plan (L'espace des configurations est de dimension 5 mais la distribution non-intégrable des déplacements infinitésimaux a dimension 2). Cette situation évoque elle-même le transport parallèle sur une sphère, encore souvent abordée en Master. L'étude des variétés roulant l'une sur l'autre est un sujet contemporain d'étude.

Le problème évoque également la notion de connexion et la théorie de l'holonomie. Certaines des approches proposées dans le concours peuvent d'ailleurs faire songer au théorème d'Ambrose et Singer.

Nous avons reçu 31 propositions pour ce problème. Seulement une dizaine d'entre elles ont été jugées hors-sujet. Les questions 1, 3 et 4 ont globalement été mieux réussies que la question 2, la plus difficile, qui concernait l'octaèdre basculant sur le plan. Finalement une douzaine de solutions ont été jugées satisfaisantes à très satisfaisantes. Toutes ces douzes sont des solutions complètes. Notons finalement que de nombreuses idées originales ou élégantes ont été proposées, pas seulement dans les meilleures copies. Cependant ces belles idées n'ont pas toujours abouti à des démonstrations solides.

Nous avons apprécié que les candidats accompagnent leurs démonstrations de figures représentant tantôt le plan, tantôt les solides. Cela semblait tout à fait indiqué dans ce problème de géométrie et a été souvent réalisé à l'aide de logiciels permettant un très beau rendu. Toutefois les figures ont parfois eu tendance à se substituer aux démonstrations. Ce fut le cas notamment dans les démonstrations faisant apparaitre un coloriage des sommets des triangles pavant le plan. Ici, en toute rigueur il était souhaitable d'indiquer que le coloriage était périodique (permettant une définition du coloriage à partir de la figure) et que des vérifications en nombre fini (on pouvait donner ce nombre) suffisaient à démontrer le résultat voulu. Une autre façon consistait à définir exactement les couleurs à l'aide de formules et d'effectuer les vérifications nécessaire sur les formules.

Compte-rendu des différentes approches. Nous listons, sans prétendre à l'exhaustivité, les approches proposées par les équipes. Précisons que ces différentes tentatives n'ont pas toujours été couronnées de succès.

Question 1 – le tétraèdre. La situation a très souvent été comprise mais pas toujours bien formulée. Voici certaines des stratégies suivies.

1. Coloriage des dalles du plan en quatre couleurs. Le fait qu'une dalle donnée ait pour voisines des dalles d'une couleur déterminée implique que l'orientation du tétraèdre est fixée parmi les trois angles a priori possibles.
2. Coloriage des sommets du plan, également en quatre couleurs (ou lettres). Variante : étude de l'orbite d'un sommet marqué du tétraèdre.
3. Pavage du plan par des patrons de tétraèdre. Cette approche n'a pas donné lieu à des solutions entièrement satisfaisantes.
4. Étude du groupe d'isométries engendré par les pivots (six basculements consécutifs autour d'un sommet du plan ramenant à la même dalle) en démontrant (de manière plus ou moins convaincante) qu'il s'agissait de tout le groupe.
5. Étude du groupe d'isométries engendré par les basculements en nombre pair. Après deux basculements le tétraèdre est globalement translaté (sans compter sa rotation propre) alors qu'après un basculement sa base subit une rotation d'angle π. Le groupe des déplacements obtenus sur le plan avec un nombre pair de basculements est isomorphe à $\mathbb{Z}^2$, un groupe plus simple que le graphe hexagonal (le dual du pavage triangulaire) qui correspond à considérer tous les basculements.
6. Démonstration de l'unicité de la position du tétraèdre sur des régions du plan de plus en plus grande (c'est une sorte de récurence). On agrandit lesdites régions en ajoutant des hexagones (six triangles) dont au moins deux appartiennent à la région précédente.
7. Approche semi-informatique, semi-exhaustive.

Question 2 – l'octaèdre. On retrouve principalement les mêmes approches.

1. Les coloriages de faces (faisant apparaitre une partition des dalles et faces en deux classes),
2. Les coloriage de sommets. Ici il convenait de colorier les sommets opposés de l'octaèdre de la même couleur. Alternativement, chaque sommet du plan se voyait attribuer deux étiquettes.
3. Nouveau par rapport au tétraèdre : coloriage des arêtes.
4. Pavages.
5. Étude du groupe d'isométries engendré par les pivots.
6. Étude du groupe d'isométries engendré par les basculements en nombre pair. Étude de son groupe dérivé.
7. Approche informatique. Au moins une équipe a proposé une borne sur le nombre de basculements suffisant à obtenir une des configurations pour une dalle dans une région limitée par périodicité.

Question 1 – l'icosaèdre. Voici ce qui a été proposé.

1. À partir d'une connaissance détaillée du groupe d'isométrie de l'icosaèdre on montre que le groupe d'isométries engendré par les pivots est ce groupe en entier.
2. On peut se placer sur n'importe quelle dalle. La dalle étant fixée, on peut remplacer la face par une face voisine à l'aide d'un pivot puis de proche en proche atteindre la bonne face. (Peu d'équipes ont pensé à indiquer que le choix du voisin peut être dirigé et non subi.) À l'aide de deux pivots, on peut changer l'orientation du contact entre la face et la salle.
3. Variante : On roule d'abord l'icosèdre de façon à obtenir la bonne face en contact avec le plan. Ensuite à partir de 5/6ème de pivot, on peut déplacer cette face sur une dalle voisine et, de proche en proche, sur n'importe quelle dalle. La fin est analogue à celle proposée précédemment.

Question 4 – le tétraèdre roulant sur l'icosaèdre. La plupart des équipes ont indiqué qu'on pouvait procéder comme pour l'icosaèdre sur le plan. D'autres stratégies ont également été proposées.

1. On peut de façon équivalente considérer que l'icosaèdre roule sur le tétraèdre. Le patron du tétraèdre pave le plan si bien qu'on peut se ramener à la question précédente (réponse pas tout à fait aboutie).
2. Le parcours exact sur comment remplacer une face par sa face opposée est détaillé. On explique alors que toute face peut être remplacée par une face distante de deux pas au plus. Ainsi les faces de chacun des deux 'hémisphères' de l'icosaèdre sont atteignables.
3. Approche informatique.

Bonus. Certaines équipe, soulignant une analogie avec la question précédente ont proposé des généralisations. En voici une compilation.

1. On peut considérer une carte planaire régulière. Chaque dalle est un polygone avec le même nombre de faces n et chaque sommet est partagé par m dalles. On peut procéder de la même façon pour les solides, avec $n' = n$ et m' faces autour de chaque sommet. Les cas dégénérés où $n' = 2$ sont évoqués. Certains candidats ont prétendu que si m et m' sont premiers entre eux toute les situations sont atteignables. C'est en fait faux comme on peut s'en rendre compte avec le cube roulant sur une grille à mailles carrées ($n' = n = 4$, $m = 4$ et $m' = 3$).

Solution

0. *Remarques préalables :*
   — La position d'un solide sur le plan est repérée par trois données : *la dalle* sur laquelle il repose, *la face en contact* et *l'orientation angulaire* des deux triangles l'un par rapport à l'autre (il y a trois angles possibles).
   — On remarque dans un premier temps qu'il est toujours possible d'amener le solide sur la dalle voulue. En effet le graphe des dalles (le graphe dual) est connexe.
   — Nous sommes amenés à nous demander : *Comment le solide peut-il tourner sur lui-même lorsque qu'une série de basculements partant de A nous ramène à cette même dalle A ?* En effet tout chemin c' de l'origine O à A se décompose comme un chemin de référence c, concaténé avec c^{-1} puis avec c', si bien qu'on se ramène à étudier l'effet des boucles $b = c' \circ c^{-1}$ entre A et A. De plus quitte à faire apparaitre $c \circ c^{-1}$ tout parcours $b \circ c$ s'écrit sous la forme $(b \circ c \circ c^{-1}) \circ c$ et donc toute boucle b sous la forme $c' \circ c^{-1}$ où c' relie 0 à A.
   — Nous allons démontrer que dans le cas 1 (tétraèdre) le solide revient toujours sur la même face avec la même orientation angulaire, dans le cas 3 et 4 (icosaèdre et tétraèdre sur l'icosaèdre) toutes les orientations sont possibles, dans le cas 2, exactement 4 des 24 orientations envisageables sont possibles. Plus précisément, par dalle, 4 faces sont possibles avec une seule orientation angulaire pour chacune. Un sous-groupe (le groupe de Klein) du sous-groupe des isométries de l'octaèdre agit simplement transitivement sur ces positions.
   — Une série de basculements spécialement intéressante est la boucle qui consiste à tourner le solide autour d'un sommet p du plan. Lors de cette opération, que nous appelons *pivot*, un des sommets du solide est constamment en contact avec le sommet p. À cet égard il est spécialement important de souligner qu'un sommet du tétraèdre est à la jonction de 3 faces (délimitées par 3 arêtes). Ce sont 4 et 5 faces pour l'octaèdre et l'icosaèdre respectivement. Un sommet du pavage est partagé par 6 dalles, ce qui implique qu'un pivot consiste en 6 basculements.
   — On peut paramètrer le réseau à l'aide des vecteurs $u = (1, 0)$ et $v = (\cos(\pi/3), \sin(\pi/3))$. On notera ainsi l'ensemble des sommets

$$R = \mathbb{Z}u + \mathbb{Z}v = \{ku + lv \in \mathbb{Z}^2 | (k, l) \in \mathbb{Z}^2\}.$$

   On a également $R = \mathbb{Z}v + \mathbb{Z}w$ et $R = \mathbb{Z}w + \mathbb{Z}u$ pour $w = v - u = (\cos(2\pi/3), \sin(2\pi/3))$.

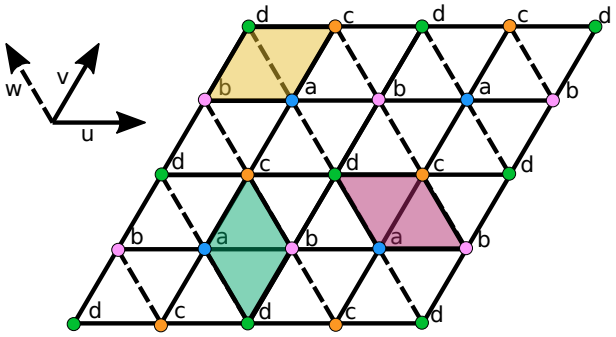

Figure 1 – Chaque losange unitaire comprend les quatre étiquettes

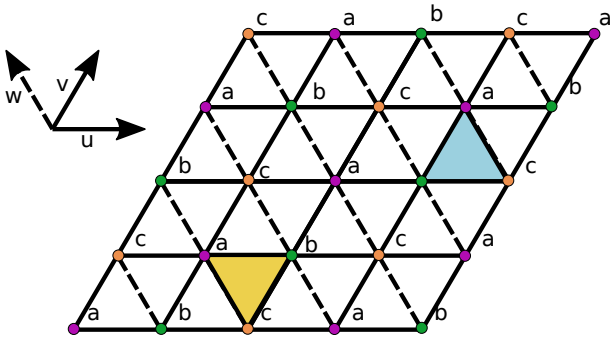

Figure 2 – Chaque triangle unitaire comprend les trois étiquettes

1. *Cas du tétraèdre.* Comme sur la Figure 1 on peut étiqueter les sommets de R à l'aide de 4 lettres (a, b, c et d) de façon à ce que *chaque losange de côté de longueur 1 (qu'on nommera* unitaire*), est marqué d'une des quatre étiquettes à ses sommets.* Techniquement cela peut-être réalisé en considérant les classes d'équivalence de $R/(2R)$. Les quatre classes sont données par l'image de l'application

$$\phi : \lambda u + \mu v \in R \simeq \mathbb{Z}^2 \mapsto (\lambda, \mu)\,[\mathrm{mod}\ 2] \in (\mathbb{Z}/2\mathbb{Z})^2$$

qui s'écrit aussi $\phi(\alpha v+\beta w) = (-\beta, \alpha-\beta)\,[2]$ ou $\phi(\gamma w+\delta u) = (\delta-\gamma, \gamma)\,[2]$. Les quatre sommets des losanges $\{\varepsilon u + \varepsilon' v\}$, $\{\varepsilon v + \varepsilon' w\}$ et $\{\varepsilon w + \varepsilon' u\}$ avec $\varepsilon, \varepsilon' \in \{0, 1\}$ portent donc des étiquettes différentes. En translatant ces trois losanges de base on obtient la même propriété pour tout les losanges unitaires.

Étiquetons le tétraèdre de façon à ce que les étiquettes des sommets en contact sur la dalle initiale coïncident avec les lettres du réseau. On s'aperçoit alors aisément que cette propriété de coïncidence est préservée après chaque basculement car les sommets impliqués dans un basculement forment un losange de côté 1. Cela nous conduit au résultat.

2. *Cas de l'octaèdre.* On colorie le plan en noir et blanc de façon à ce que les dalles voisines soit de couleurs différentes. Cela est également réalisable pour l'octaèdre (qui peut être vu comme la jonction le long d'un équateur deux de deux pyramides à base carrées). On fait de sorte à ce que dans la position d'origine les couleurs coïncident. Cette coïncidence étant préservée par les basculement, il s'avère qu'au maximum quatre faces de l'octaèdre peuvent « établir le contact » avec une dalle donnée. Cependant chacune de ces faces peut y parvenir comme cela se montre à l'aide des pivots. Chacun des trois pivots permet en effet, dans un sens ou dans l'autre, de remplacer une face par l'unique face de même couleur qui partage avec elle le sommet de l'octaèdre utilisé pour pivoter (c'est le sommet qui au cours de la manœvre est constamment en contact avec le plan).

Finalement en étiquetant (= coloriant) R par les trois classes de $S = \mathbb{Z}(u+v) + \mathbb{Z}(2u-v) = \mathbb{Z}(3u) + \mathbb{Z}(u+v)$ ainsi que l'octaèdere par les trois classes d'équivalence de la symétrie centrale, on démontre que l'orientation angulaire est prescrite par le couple dalle/face (cf. Figure 2). Avant de voir cela en détail, notons tout d'abord que S est le noyau du morphisme de groupe

$$\psi : \lambda u + \mu v \in R \simeq \mathbb{Z}^2 \mapsto u - v\,[3] \in \mathbb{Z}/3\mathbb{Z}.$$

Ainsi y a-t-il trois classes d'équivalence . De plus u, v, w n'étant pas dans le noyaux de ψ, les triangles équilatéraux ont des sommets des trois couleurs.

Le principe est le même qu'au 1. Lors d'un basculement deux sommets de l'octaèdre conservent le contact avec le plan. Le troisième est remplacé par le sommet de même couleur opposé sur l'octaèdre. Les triangles équilatéraux de R font apparaitre les trois étiquettes si bien que le sommet du plan nouvellement impliqué dans le contact avec l'octaèdre est de même couleur que celui qui perd cette propriété. La propriété de coïncidence des couleurs se transmet donc héréditairement au cours d'un basculement.

3. *Cas de l'icosaèdre.* Le cas 4 est analogue au 3. Nous nous contenterons de la correction du 3. Rappelons que l'on peut toujours atteindre une dalle donnée. Sur cette dalle les différents pivots envisageables permettent de remplacer une face par chacune de ses trois voisines. Une série de pivots permet donc de conserver la dalle et d'atteindre la face voulue, par connexité du graphe des faces de l'icosaèdre. Finalement, toutes les orientations angulaires sont possibles : pour obtenir un tiers de tour on procède à deux pivots consécutifs de sommets différents, le premier changeant la face, le second rétablissant la première face.

Nicolas Juillet, Université de Strasbourg

Problème 5 : Modélisation

Auteur : Filippo Santambrogio (Université Paris-Sud)

Une personne se trouve sur le trottoir sud d'une longue route rectiligne orientée est-ouest, et doit la traverser en se déplaçant en même temps d'une certaine distance (fixée, et non nulle) vers l'ouest.

Elle marche à une vitesse prescrite et constante et voudrait arriver à destination le plus rapidement possible. Cependant, tant qu'elle marche en pleine chaussée, il y a un risque de se faire renverser par une voiture, et ce risque (par unité de temps) est une fonction croissante de la distance aux trottoirs (donc, maximal au milieu de la chaussée, et croissant en s'y rapprochant ; on suppose aussi que cette fonction "risque" est une fonction régulière de la position où on se trouve). Elle décide donc de choisir sa trajectoire de manière à minimiser la somme du temps et d'un coût proportionnel au risque total de se faire renverser pendant le parcours.

Prouver qu'une telle trajectoire optimale existe, et en discuter les propriétés qualitatives :

— S'agit-il du graphe d'une fonction ?

— Quelle est l'équation différentielle satisfaite par la courbe et/ou le graphe ?

— Que peut-on dire de la courbure de cette trajectoire (concavité ou convexité du graphe) ?

Prouver également l'unicité de cette trajectoire optimale.

Rapport

Le problème portait sur une question très naturelle et issue de la vie réelle : comment traverser une route lorsqu'on doit en même temps se déplacer en direction parallèle à la route elle même. La plupart des enfants ont envie de traverser tout droit en diagonale pour arriver le plus vite possible, alors que leurs mères, soucieuses de ne pas les laisser trop longtemps au milieu de la chaussée où ils pourraient se faire renverser par une voiture, voudraient qu'ils fassent un chemin en L en restant le plus possible sur le trottoir. . .. Où est le juste milieu ?

Ce genre de questions, qui reviennent à chercher une géodésique pour une longueur pondérée avec un coût pour passer par chaque point qui est non-uniforme, reviennent souvent en recherche. Souvent l'espace parcouru par la courbe que l'on cherche n'est pas le plan (comme dans cet exercice) mais un espace de fonctions (récemment cela a été utilisé pour résoudre des problèmes d'EDP avec manque de compacité), de mesures (en transport optimal, mécanique des fluides. . .). . . ou alors il s'agit bien d'une trajectoire dans l'espace euclidien d'un individu, mais le modèle demande ensuite à en faire interagir plusieurs (comme dans les jeux à champ moyen. . .).

Les réponses à cette question de modélisation ont été très variées, et on y trouve plusieurs bonnes idées ainsi que plusieurs erreurs ou anomalies. La première chose à faire consistait à traduire le problème, décrit exprès de manière très vague dans l'énoncé, en une question mathématique. À quelques exceptions près cela a été bien fait : on peut néanmoins distinguer les étudiants qui ont décrit la trajectoire γ du piéton comme une coube lipschitzienne à vitesse constante et prescrite, ceux qui ont imposé une vitesse maximale (qui sera alors saturée à l'optimum), et ceux qui ont laissé la vitesse complètement libre mais mis un facteur $|\gamma'|$ dans l'intégrale (ce à quoi on peut arriver après un changement de variable). Ceux qui ont imposé $|\gamma'| = const$ ont souvent oublié de prendre en compte le fait qu'une limite (uniforme) de courbes γ_n satisfaisant cette condition ne la satisfait pas forcément ; parmi ceux qui ont utilisé le facteur $|\gamma'|$, il y en a qui l'ont mis seulement sur une partie du coût à minimiser (le temps de parcours et non la partie risque). Presque personne n'a traduit le problème en une minimisation d'énergie avec $|\gamma'|^2$, qui simplifie beaucoup les arguments d'existence et l'écriture d'une équation d'Euler-Lagrange.

Si des étudiants ont en effet prouvé l'existence par des arguments basés sur la méthode directe (prendre une suite minimisante, en extraire une suite qui converge, prouver que la limite minimise. . ., ce qui a posé des difficultés à certains qui ont affirmé que les quantités qu'on minimise, essentiellement la longueur, sont continues pour la convergence uniforme, alors qu'elles ne sont que semi-continues inférieurement, mais cela suffit pour l'existence d'un minimum), la plupart de ceux qui ont touché à cette question l'ont fait en évoquant le Théorème d'Hopf-Rinow et d'autres notions de géométrie Riemannienne. Ceci est correct, mais un peu décevant, alors qu'une preuve ad hoc pouvait être faite de manière assez élémentaire.

Au contraire, en ce qui concerne la preuve du fait que la courbe soit un graphe ou l'étude de sa convexité, les étudiants ont surtout utilisé des arguments "à la main" (remplacer une portion de courbe par un segment et montrer que le segment a un coût inférieur, ce qui soulevait parfois des difficultés que les étudiants n'ont pas toujours vu), alors qu'il était possible d'utiliser davantage l'équation d'Euler-Lagrange. Ceci permettait de dire rapidement que la courbe est un graphe sur la variable x (abscisse parallèle à la direction de la route) par le fait qu'on obtenait $x' > 0$. Le fait que la courbe soit aussi un graphe par rapport à l'autre variable était plus compliqué.

En effet, il y avait une subtilité, pas forcément voulue, dans l'interprétation du sujet (les hypothèses étaient, exprès, très vagues). Dans le sujet on parlait d'une fonction "risque", nulle sur les trottoirs et régulière : on peut la considérer comme définie partout, et donc sa dérivée doit s'annuler sur le bord du trottoir, ou définie seulement sur la chaussée, ce qui permet une dérivée non nulle sur le bord. La plupart des théorèmes de géométrie évoqués par les étudiants concernent des variétés sans bord et lisses. Si on se place donc dans le premier cas tout marche bien, et on a aussi le résultat d'unicité demandé ainsi que le fait que la courbe soit un graphe par rapport à y.

Trois groupes d'étudiant ont bien compris la subtilité (eux aussi, ils n'en sont pas forcément conscients) : un a bien clarifié qu'il utilise une dérivée nulle sur le bord comme conséquence du fait que la fonction est nulle sur le trottoir et lisse ; deux autres ne font pas cette hypothèse mais

font la distinction dans les résultats d'unicité d'après la dérivée au bord. Ces deux copies trouvent aussi que, dans le cas de dérivée non nulle au bord, l'unicité est vérifiée si le déplacement vers l'est est suffisamment petit. Une de ces copies, par contre, ne prouve pas l'existence, alors que l'autre le fait par des méthodes directes, ce qui est remarquable. C'est cette copie qui a eu la note la plus élevée, sans pour autant qu'il s'agisse d'un 10/10, car certaines constructions géométriques qui sont faites pour améliorer les courbes ne sont pas 100% justifiées.

Solution

Il faut d'abord transformer cet énoncé en un problème de mathématiques. Représentons la route comme le sous-ensemble $R = \{(x,y) \in \mathbb{R}^2 : |y| \leq 1\}$ du plan $\mathbb{R}^2$ et les points de départ de d'arrivée comme étant $P = (-L,-1)$ et $Q(L,1)$, respectivement (donc $2L > 0$ est le déplacement prescrit vers l'ouest). La trajectoire parcourue sera représentée par une courbe $\gamma = [0,T] \to R$ avec $\gamma(0) = P$ et $\gamma(T) = Q$. La valeur de T n'est pas fixée. L'information "elle marche à une vitesse prescrite et constante" nous permet de dire que γ est une courbe lipschitzienne (parce que si la vitesse est constante, alors elle est bornée), avec $|\gamma'(t)| = V$ p.p. (les fonctions lipschitziennes étant différentiables p.p.). Le but est de minimiser la quantité

$$J(T,\gamma) := T + \int_0^T c(\gamma(t))dt,$$

et plus précisément de résoudre

$$\min\{J(T,\gamma) \ : \ \gamma \in \mathrm{Lip}([0,T];R), \gamma(0) = P, \gamma(T) = Q, |\gamma'| = V\},$$

où c représente le "risque de se faire renverser". Attention : l'énoncé du sujet demandait à faire la mathématisation, donc des choix s'imposent : on peut considérer c comme étant défini sur R et régulier sur R, ou c comme étant défini sur R^2, nul en dehors de R, et régulier sur $\mathbb{R}^2$ (donc en particulier ses dérivées s'annulent sur ∂R). Dans ce cas, on devrait imposer que les courbes γ soient à valeurs dans R, mais cela n'est pas important, il est facile de voir que comme c est constant en dehors de l'intérieur de R il n'est pas intéressant de sortir. C'est ce dernier point de vue (c régulier sur $\mathbb{R}^2$) qu'on considère ici.

Existence. On prend une suite minimisante : des courbes γ_n, chacune définie sur un intervalle $[0,T_n]$, telle que $J(T_n,\gamma_n) \to \ell$, la valeur ℓ étant celle de l'infimum de notre problème. Comme $J(T_n,\gamma_n)$ est borné supérieurement, T_n est borné aussi et on peut supposer, à une sous-suite près, que l'on ait $T_n \to T$. Prenons $T' < T$: à partir d'un certain rang, on aura $T_n > T'$ et donc les courbes γ_n sont définies sur $[0,T']$ et y sont V-lipschitziennes. On peut donc extraire une sous-suite qui converge uniformément sur $[0,T']$ (grâce au théorème d'Ascoli-Arzelà) et, en faisant cela pour des temps T' qui approchent T et en prenant une sous-suite diagonale, on peut supposer que la même sous-suite γ_n converge uniformément sur tous les intervalles de type $[0,T']$ pour $T' < T$ vers une même courbe limite $\gamma : [0,T] \to R$. Cette courbe sera V-lipschitzienne sur chaque intervalle $[0,T']$, ce qui peut se prouver en passant à la limite la propriété $|\gamma_n(t) - \gamma_n(s)| \leq V|t-s|$ pour $t,s \in [0,T']$. De plus, on a

$$|\gamma_n(t) - Q| = |\gamma_n(t) - \gamma_n(T_n)| \leq V|t - T_n|$$

et, en passant à la limite pour $n \to \infty$, on obtient, pour $t < T$, l'estimation $|\gamma(t) - Q| \leq V|t-T|$, qui montre que la courbe γ peut être prolongée par continuité en $t = T$ avec $\gamma(T) = Q$.

En utilisant $T_n \to T$ et $\int_0^{T'} c(\gamma_n(t))dt \to \int_0^{T'} c(\gamma(t))dt$ (par convergence uniforme), on obtient

$$\int_0^{T'} c(\gamma(t))dt \leq \liminf_n \int_0^{T_n} c(\gamma_n(t))dt = \ell$$

et donc $J(T,\gamma) \le \liminf_n J(T_n,\gamma_n)$. Cela démontrerait que γ est optimale si elle était admissible : on doit montrer $\gamma(0) = P$ (ce qui s'obtient de $\gamma_n(0) = P$ par limite uniforme), $\gamma(T) = Q$ (ce qu'on a déjà justifié) et $|\gamma'| = V$. Pour l'instant on sait juste que γ est V-lipschitzienne, donc $|\gamma'| \le V$. Supposons qu'on n'a pas $|\gamma'| = V$ p.p. : on peut alors trouver un reparamétrage de γ, qu'on appelle $\tilde\gamma$, défini sur un intervalle $[0,\tilde T]$ avec $\tilde T < T$, et ayant une vitesse constante égale à V : $|\tilde\gamma'| = V$. Or, cette courbe est sûrement admissible, et donc on a $J(\tilde T,\tilde\gamma) \ge \ell$, ainsi que

$$\begin{aligned}\ell &\ge J(T,\gamma) := T + \int_0^T c(\gamma(t))dt > \tilde T + \int_0^T c(\gamma(t))\frac{|\gamma'|}{V}dt \\ &= \tilde T + \int_0^{\tilde T} c(\tilde\gamma(t))\frac{|\tilde\gamma'|}{V}dt = \tilde T + \int_0^{\tilde T} c(\tilde\gamma(t))dt = J(\tilde T,\tilde\gamma) \ge \ell,\end{aligned}$$

ce qui est une contradiction. On a utilisé l'inégalité $|\gamma'|/V \le 1$ et l'invariance par réparamétrage des intégrales du type $\int a(\gamma)|\gamma'|$, et on a démontré l'existence d'une courbe optimale.

Attention. Des théorèmes généraux de géométrie riemannienne (le théorème de Hopf et Rinow) peuvent aussi permettre d'établir l'existence d'une courbe optimale, en tant que géodésique pour la métrique conforme induite par $1 + c$, sous hypothèses de régularité sur c.

Reformulation. On utilise l'invariance par réparamétrage pour réduire le problème à un problème sur un intervalle fixé, et sans contraintes de vitesse. Tout d'abord on écrit

$$J(T,\Gamma) = \int_0^T c_1(\gamma(t))dt = \int_0^T \bar c(\gamma(t))\frac{|\gamma'(t)|}{V}\,dt = \int c_2(\gamma(t))|\gamma'(t)|dt,$$

où $c_1 = 1+c$ et $c_2 = c_1/V$. Cette dernière expression montre que J peut s'écrire comme une quantité qui ne dépend pas du paramétrage. Comme toute courbe lipschitzienne peut se reparamétrer à vitesse constante égale à V, le problème peut se reformuler comme

$$\min\left\{\int_0^1 c_2(\gamma)|\gamma'|\,dt \;:\; \gamma\in \mathrm{Lip}([0,1];R),\ \gamma(0)=P,\gamma(1)=Q\right\},$$

où l'intervalle $[0,1]$ a été choisi de manière arbitraire. La valeur du minimum vaut toujours ℓ et le minimum ne sera pas réalisé par une unique courbe γ, parce que tous ses reparamétrages donneront la même valeur.

Reste la difficulté que, pour écrire des conditions d'optimalité, on a envie de dériver, et la norme $|\gamma'|$ n'est pas une fonction dérivable. On peut alors remarquer que l'on a

$$H(\gamma) := \int_0^1 c_2^2(\gamma)|\gamma'|^2\,dt \ge \left(\int_0^1 c_2(\gamma)|\gamma'|\,dt\right)^2 \ge \ell^2,$$

où la première inégalité est justifiée par l'inégalité de Jensen (et on a égalité si et seulement si $c_2(\gamma)|\gamma'|$ est constante), et la deuxième est la définition de ℓ (et on a égalité si et seulement si γ est optimal) De fait, cela signifie qu'on peut minimiser H, et que cela ne fera que sélectionner les minimiseurs qui satisfont la condition $c_2(\gamma)|\gamma'|=$ constante. On considère donc une fonctionnelle du type $H(\gamma) = \int_0^1 L(\gamma(t),\gamma'(t))\,dt$ où $L(x;v) := a(x)|v|^2$ et $a = c_2^2$.

Conditions nécessaires d'optimalité On utilise l'*équation d'Euler-Lagrange* satisfaite par les minimiseurs des fonctionnelles $H(\gamma) := \int_0^1 L(\gamma,\gamma')$ où $L : \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ est C^1 et la minimisation est faite parmi les courbes avec $\gamma(0)$ et $\gamma(1)$ fixés. Cette équation, obtenue en imposant

$$\frac{d}{d\varepsilon}\left(H(\gamma+\varepsilon\eta)\right)_{|\varepsilon=0} = 0$$

pour une perturbation $\eta \in C_c^\infty(]0,1[)$ arbitraire, nous donne

$$\frac{d}{dt}\left(\nabla_v L(\gamma(t),\gamma'(t))\right) = \nabla_x L(\gamma(t),\gamma'(t)).$$

Dans notre cas, on obtient $(2\gamma'(t)a(\gamma(t)))' = |\gamma'(t)|^2\nabla a(\gamma(t))$. En écrivant $\gamma(t) = (x(t), y(t))$ et en utilisant le fait que a ne dépend que de y (et on écrira donc $a(y)$ au lieu de $a(\gamma)$) on peut écrire cela comme un système

$$\begin{cases} (a(y(t))x'(t))' = 0, \\ (a(y(t))y'(t))' = (|x'(t)|^2 + |y'(t)|^2)a'(y(t)). \end{cases} \tag{29}$$

Attention : ces équations ne sont plus valables dans le cas où c est défini sur R et γ contrainte à rester dans R, du fait de cette même contrainte (elle serait à remplacer par une inégalité sur le bord de R, c'est-à-dire où $y = \pm 1$; cependant, on peut vérifier qu'elles restent vraies dans le cas où $a'(\pm 1) = 0$).

Propriétés de la solution Le système (29) répond à la question "quelle équation différentielle satisfait la courbe optimale ?". Mais on peut dire plus. En utilisant la première ligne de (29) on a $a(y(t))x'(t) = C$. Comme x' ne peut pas être partout nulle, et que $a > 0$, on déduit que x' ne s'annule jamais. Ceci permet de dire que la courbe $(x'(t), y'(t))$ peut s'écrire comme un graphe $y = f(x)$. On peut écrire l'équation différentielle qui régit f : en dérivant $y(t) = f(x(t))$ on obtient

$$y'(t) = f'(x(t))x'(t) = f'(x(t))\frac{C}{a(y(t))}$$

donc $a(y(t)y'(t) = Cf'(x(t))$. Si on dérive et on utilise la deuxième ligne de (29) on obtient

$$\begin{aligned} \frac{C^2}{a^2(y(t))}(1 + |f'(x(t))|^2)a'(y(t)) &= (|x'(t)|^2 + |y'(t)|^2)a'(y(t)) = ((a(y(t))y'(t))' \\ &= C\,(f'(x(t)))' = f''(x(t))\frac{C^2}{a(y(t))}. \end{aligned}$$

Ceci donne

$$f''(x) = \frac{a'(f(x))}{a(f(x))}(1 + |f'(x)|^2), \tag{30}$$

qui est l'équation différentielle satisfaite par f. Ceci permet tout d'abord d'établir la régularité de f : si $a \in C^\infty$, on peut démontrer par récurrence que $f \in C^k$ pour tout k. Aussi, on peut déduire la concavité/convexité de f : f est convexe là où $a' > 0$ et concave là où $a' < 0$. Grâce aux hypothèses (que a est une fonction croissante de la distance aux trottoirs) on déduit que f est convexe jusqu'à $y = 0$ et concave ensuite.
Attention : il est également possible de prouver que la fonction est un graphe par rapport à la variable y, mais moins évident. Un preuve possible est de remplacer toute partie de courbe entre deux points avec la même ordonnée par un segment, mais cela ne réduit pas toujours le coût ; pour conclure, il faut aussi "réordonner" les parties de la courbe, en mettant le segment à la fin, sur le bord de R. Par contre, l'équation sur f^{-1} est beaucoup plus simple est peut être résolue en calculant une primitive.

Unicité On cherche à prouver l'unicité de la fonction f dont on considère le graphe. Or, l'équation (30) est une équation ordinaire dont on a existence et unicité du problème de Cauchy. Considérons un point x_0 où on a $f(x_0) = 0$ (il y en a un car $f(\pm L) = \pm 1$). On ne connaît pas $f'(x_0)$ mais la solution de (30) doit être symétrique par rapport au point $(x_0, 0)$, par unicité du problème de Cauchy (en utilisant la parité de la fonction a). Cela signifie que f touche les valeurs $y = \pm 1$ à une même distance de x_0. Il est clair que la courbe optimale ne peut pas toucher $y = \pm 1$, s'en éloigner et y revenir (ce ne serait pas optimal, il vaudrait mieux rester sur l'axe horizontal $y = \pm 1$, ce qui coûterait moins cher parce que a est minimale en ± 1), et on ne peut pas avoir de parties plates constantes égale à ± 1 dans le graphe de f (sinon on aurait une région avec $f'' = f' = 0$ et, par unicité de la solution de (30), la courbe coïnciderait avec la constante ± 1, ce qui n'est pas le cas). Cela prouve $x_0 = 0$ et la symétrie de f. Le coût total de cette courbe est donc reparti de manière

égale entre la partie $x < 0$ (qui correspond aussi à $y < 0$) et $x > 0$ ($y > 0$). On n'a pas encore l'unicité de f parce qu'on pourrait avoir deux graphes différents, avec deux valeurs différentes pour $f'(0)$. Cependant, si tel était le cas, on pourrait prendre deux de ces graphes et les coller en $x = 0$, en obtenant une nouvelle courbe, elle aussi optimale (parce que toutes les courbes optimales ont le même coût sur les deux intervalles $[-L, 0]$ et $[0, L]$). Mais cette courbe correspondrait à un graphe qui devrait aussi résoudre (30), mais qui ne serait pas C^1, ce qui donne une contradiction !
Attention : dans le cas où c est défini sur R et que ses dérivées ne s'annulent pas sur le bord de R on ne peut pas obtenir la même conclusion : on a unicité de la forme de la courbe optimale à l'intérieur de R, mais elle peut contenir des segments sur le bord, et peut donc être translatée. On a unicité pour L petit (auquel cas on n'a pas de partie plate) mais pas pour L grand (et la solution est un graphe sur x mais pas sur y).

Filippo Santambrogio, Université Paris-Sud

Problème 6 : Probabilités

Auteur : Olivier Garet (Université de Lorraine)

Soit μ une loi sur $\mathbb{Z}$ fixée. Pour $k \in \mathbb{Z}$, on appelle *marche aléatoire* partant de k et de loi des pas μ une suite de variables aléatoires $(S_n)_{n\geq 0}$ définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ telle que

- $S_0 = k$ presque sûrement,
- la suite des pas $(X_n)_{n\geq 1}$, définie pour tout $n \geq 0$ par

$$X_{n+1} = S_{n+1} - S_n, \tag{31}$$

est une suite de variables aléatoires indépendantes de loi μ.

Dans la suite, on notera $\mathbb{P}^k$ une mesure de probabilité sur $(\Omega, \mathcal{F})$ telle que sous $\mathbb{P}^k$, $(S_n)_{n\geq 0}$ est une marche aléatoire partant de k et de loi des pas μ.

Le but du problème est de décrire une méthode permettant de déterminer la suite $(u_k)_{k\in\mathbb{Z}}$, définie par

$$\forall k \in \mathbb{Z} \quad u_k = \mathbb{P}^k(\forall n \geq 0,\ S_n \geq 0). \tag{32}$$

1. On se propose d'étudier d'abord le cas particulier suivant :

$$\mu = \frac{3}{4}\delta_1 + \frac{1}{4}\delta_{(-2)}. \tag{33}$$

Montrer que la suite $(u_k)_{k\in\mathbb{Z}}$ [3] satisfait une relation de récurrence linéaire que l'on déterminera.
Montrer que $\lim_{k\to+\infty} u_k = 1$, puis exprimer u_k en fonction de k.

2. On suppose maintenant que la loi μ est à support dans un ensemble fini $\{-p, \dots, q\} \subset \mathbb{Z}$ et telle que $\mathbb{E}(X_1) > 0$.
Déterminer une méthode, la plus générale possible, permettant d'exprimer u_k en fonction de k.

3. L'énoncé initial mentionnait $(u_k)_{k\in\mathbb{Z}}$, mais cette coquille n'a gêné personne.

Rapport

Une question fréquemment posée à propos d'une chaîne de Markov, c'est de déterminer avec quelle probabilité un événement va arriver avant un autre, ou s'il va arriver un jour.

Ces problèmes peuvent souvent être résolus à partir de la méthode de l'analyse au premier pas : le point de départ étant considéré comme un paramètre, en conditionnant par la valeur du premier mouvement de la chaîne, on obtient alors un système d'équations linéaires.

Dans le cas où la chaîne de Markov est une marche aléatoire, le système donne une récurrence linéaire à coefficients constants. L'exemple le plus classique est le problème de la ruine du joueur, qui est traité dans tous les cours de M1. Ces systèmes d'équations sont souvent résolus par l'adjonction de conditions "aux limites", qui correspondent à des positions où l'issue est déterministe.

On est dans un schéma assez classique et la première question, assez guidée, ne devait pas poser de problèmes à de bons étudiants (de fait une seule copie non vide a échoué à résoudre cette question). Cependant, le succès de la méthode n'est pas assuré car rien n'assure a priori de disposer de suffisamment d'équations pour déterminer le système.

La question ici posée consistait à déterminer la probabilité qu'une marche aléatoire à support fini, biaisée dans le sens positif, ne passe pas dans les négatifs. Ici, les premières conditions au bord sont déterminées par le fait que partant de valeurs négatives, la probabilité est évidemment nulle. Avec un peu de travail, on montre que quand le point de départ s'échappe à l'infini, la probabilité tend vers 1, ce qui nous donne encore une condition.

A-t-on assez de conditions ? La démarche que j'avais envisagée, et qui a été suivie par presque toutes les copies est la suivante : on résout une équation de récurrence linéaire, on l'exprime à l'aide de la théorie générale ; u_n s'écrit

$$u_n = \sum_i P_i(n)\lambda_i^n,$$

où les λ_i sont les racines de l'équation caractéristique de la marche. Reste à déterminer les polynômes qui s'affichent en face des puissances. En face des puissances des racines de module dépassant 1, le polynôme doit être 0, sinon les probabilités explosent. Cette « presque évidence »demande tout de même un petit travail, qui a été parfois ignoré, parfois plus ou moins bien fait. Ceci étant fait, ne reste-t-il pas trop de racines ? Ce dernier point est traité avec le lemme de Rouché, que quelques-uns malmènent, et qu'une copie réinvente dans un cas particulier. Le soin donné aux preuves des points évoqués ci-dessus départage ainsi l'essentiel des copies.

Cependant, une copie sort du lot, qui utilise un argument assez général, et cette fois très probabiliste, auquel je n'avais pas songé. Je le décris brièvement. Il est facile de voir que la suite des probabilités cherchées constitue une fonction harmonique bornée de l'opérateur associé à la chaîne de Markov, ce qui permet de lui associer une martingale. Mais en réalité, toute solution bornée du système que l'on a obtenu est aussi une fonction harmonique, peut aussi se voir associer une martingale ; ainsi à l'aide d'un théorème de martingale, on peut identifier la solution et la suite cherchée, ce qui montre que le système linéaire est bien déterminé. L'argument est très pur, et visiblement généralisable. La copie n'utilise pas le vocabulaire des fonctions harmoniques. Que ses auteurs soient cultivés ou très malin reste en débat ; enfin il est clair qu'ils m'ont fait comprendre quelque chose.

Pour une référence en français pour la méthode "first step analysis", je recommande mes notes de cours

Probabilités et processus stochastiques
`http://www.iecl.univ-lorraine.fr/~Olivier.Garet/livre-pps/`

Solution

1. On cherche la probabilité qu'une marche aléatoire biaisée $\frac{3}{4}\delta_1 + \frac{1}{4}\delta_{-2}$ partant de n reste tout le temps dans $\mathbb{N}$. Si on note u_n la probabilité que la marche partant de n reste tout le temps dans $\mathbb{N}$, la propriété de Markov donne pour $n \geq 0$.

$$u_n = \frac{1}{4}u_{n-2} + \frac{3}{4}u_{n+1}.$$

L'équation polynômiale associée $3x^3-4x^2+1$ admet les racines $1, \phi = \frac{1+\sqrt{13}}{6}, \overline{\phi} = \frac{1-\sqrt{13}}{6}$, donc $u_n = A\phi^n + B\overline{\phi}^n + C$. On a les conditions au bord $u_{-1} = u_{-2} = 0$. Une dernière condition est donnée par $\lim_{n\to+\infty} u_n = 1$. La preuve en sera donnée dans le cas général, à la question 2. Admettons-le pour l'instant. Cela nous donne $C = 2$.
On a alors un système 2×2 à résoudre. Les calculs sont un peu pénibles, donc je prends giac (Xcas).
(On peut faire les calculs en ligne avec
`http://www.xcasenligne.fr/giac_online/demoGiacPhp.php`).

```
lphi:=(1+sqrt(13))/6
lphibar:=(1-sqrt(13))/6
linsolve([a*lphi^(-2)+b*(lphibar)^(-2)=-1,
a*lphi^(-1)+b*lphibar^(-1)=-1],[a,b])

[(-4*sqrt(13)-13)/39,(4*sqrt(13)-13)/39]
```

Ainsi, on a

$$u_n = 1 + \frac{-4\sqrt{13}-13}{39}\left(\frac{1+\sqrt{13}}{6}\right)^n + \frac{4\sqrt{13}-13}{39}\left(\frac{1-\sqrt{13}}{6}\right)^n,$$

ce qui nous donne en particulier $u_0 = \frac{13}{39} = \frac{1}{3}$. Comme u_{-1}, u_{-2}, u_0 sont rationnels, il est maintenant immédiat par récurrence que u_k est rationnel pour tout k. Ce n'était pas demandé, mais c'est amusant.

2. Quelques notations supplémentaires : on introduit la fonction génératrice de la marche (c'est à dire de la loi des pas) : on a

$$f(z) = \mathbb{E}[z^{X_1}] = \sum_{i=-p}^{q} a_i z^i,$$

avec les a_i positifs, de somme 1. a_{-p} et a_q sont non nuls. On pose enfin $m = \mathbb{E}(X_1)$.
Notons u_n la probabilité pour que la marche partant de n reste dans $\mathbb{Z}_+$ et $v_n = 1 - u_n$ la probabilité pour que la marche partant de n entre dans $\mathbb{Z}_-^*$. Bien sûr connaître l'un, c'est connaître l'autre.
On pose donc

$$E = \{\exists n \geq 0; S_n < 0\}$$

La suite (S_n) étant une chaîne de Markov, on peut appliquer la méthode de l'analyse au premier pas.
Pour $k \in \mathbb{N}$, les événements E et

$$\theta^{-1}(E) = \{\exists n \geq 1; S_n < 0\}$$

coïncident $\mathbb{P}^k$ presque sûrement : on a

$$\begin{aligned} v_k &= \mathbb{P}^k(E) = \mathbb{P}^k(\theta^{-1}(E)) \\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}^k(S_1 = i, \theta^{-1}(E)) \\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}^k(S_1 = i)\mathbb{P}^i(E) \\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}^k(S_1 = k+i)\mathbb{P}^{k+i}(E) \\ &= \sum_{i=-p}^{q} a_i v_{k+i} \\ &= (f(\theta).v)_i, \end{aligned}$$

où on a noté θ l'opérateur de translation : $(\theta u)_n = u_{n+1}$.
On a ainsi une équation de récurrence linéaire $(I - f)(\theta).u = 0$, ou encore $Q(\theta).u = 0$, où on a posé $Q(z) = (1 - f(z))z^p$ On sait bien que si le polynôme Q se décompose dans $\mathbb{C}[X]$ en

$$Q(X) = a_p \prod_{i\in\mathcal{I}} (X - \lambda_i)^{n_i},$$

les suites solutions de ces équations sont exactement les suites de la forme

$$v_n = \sum_{i\in\mathcal{I}} P_i(n)\lambda_i^{n_i}, \tag{34}$$

où P_i est un polynôme de degré inférieur à n_i. [4] Notons au passage que, comme $a_{-p} \neq 0$, les racines de l'équation $f(z) = 1$ sont exactement les racines du polynôme Q.
Toute la question est de savoir si nous avons suffisamment d'informations pour être capable de déterminer les P_i. On a évidemment les conditions "au bord" : pour tout $i \in \{1,\dots,p\}$, $v_{-i} = 1$. On peut également démontrer que l'on a

$$\lim_{n\to+\infty} v_n = 0.$$

Preuve 1 : En effet si S_n est la suite des sommes partielles d'une telle marche partant de 0,

$$\begin{aligned} v_n &= \mathbb{P}(\exists k \geq 0; n + S_k < 0) \\ &\leq \sum_{k\geq 0} \mathbb{P}(n + S_k < 0) \\ &\leq \sum_{k\geq 0} \mathbb{P}(\exp(-\alpha S_k) \geq \exp(-\alpha n)) \\ &\leq \exp(-\alpha n) \sum_{k\geq 0} \mathbb{E}[\exp(-\alpha S_k)] \end{aligned}$$

ceci quelque soit $\alpha > 0$. Bien sûr $\mathbb{E}[\exp(-\alpha S_k)] = (\mathbb{E}[\exp(-\alpha X_1)])^k$, et si on prend α assez petit, la série est convergente car

$$\mathbb{E}[\exp(-\alpha X_1)] = 1 - \alpha\mathbb{E}[X_1] + o(\alpha)$$

et $\mathbb{E}[X_1] = m > 0$.

4. Ceci ce démontre à l'aide du lemme des noyaux, voir par exemple Cours de Mathématiques MP/MP*, Voedts, p.195.

Preuve 2 : si on note $A_n = \{\exists k \geq 0; n + S_k < 0\}$, on peut noter que la suite A_n est décroissante, d'intersection

$$A = \cap_{n\geq 1} A_n = \{\inf_{k\geq 1} S_k = -\infty\},$$

donc avec le théorème de continuité séquentielle décroissante $\lim v_n = \mathbb{P}(A)$. Or, la loi forte des grands nombres nous dit que S_k/k converge vers $\mathbb{E}(X_1) > 0$: presque sûrement, à partir d'un certain rang, tous les S_k sont positifs : leur borne inférieure est presque sûrement finie.

De surcroit, comme la suite $(v_n)_{n\geq 0}$ est de limite nulle, les polynômes P_i correspondants aux λ_i de module supérieur ou égal à 1 doivent être nuls.

En effet, soit J l'ensemble des index qui apportent la plus grosse contribution, c'est à dire qui maximisent $(|\lambda_i|, \deg P_i)$ pour l'ordre lexicographique. On note (λ_0, d_0) ce maximum. On va montrer que si $\lambda_0 \geq 1$, on arrive à une contradiction. En effet, en divisant par $\lambda_0^n n^{d^0}$ l'équation (34), on obtient

$$\frac{v_n}{\lambda_0^n n^{d^0}} = \sum_{i\in J} c_i(\lambda_i/\lambda)^n + O(1/n),$$

où c_i est le coefficient du terme de plus haut degré de P_i. Si $\lambda_0 \geq 1$, cela nous donne

$$\sum_{i\in J} c_i(\lambda_i/\lambda)^n = o(1).$$

Fixons $i_0 \in J$: en multipliant par $(\lambda/\lambda_{i_0})^n$ (qui est de module 1), on a encore

$$\sum_{i\in J} c_i(\lambda_i/\lambda_{i_0})^n = o(1).$$

On en déduit que la suite $(\sum_{i\in J} c_i(\lambda_i/\lambda_{i_0})^n)_{n\geq 1}$ converge en moyenne de Cesaro vers 0 ; or un calcul simple montre qu'elle converge en moyenne de Cesaro vers c_{i_0}, donc $c_{i_0} = 0$: contradiction.

Cela nous donne un système d'équations linéaires. La suite vise à montrer que ce système est suffisant pour déterminer $(v_n)_{n\geq -p}$ (ou $(u_n)_{n\geq -p}$).

Ainsi $(v_n)_{n\geq -p}$ appartient à l'espace vectoriel E des solutions de

$$(1 - f)(\theta).u = 0$$

correspondant aux racines de module strictement inférieur à 1. Notons I l'ensemble de tels index.

Nous allons maintenant montrer que l'équation $f(z) = 1$ a exactement p racines (en comptant les ordres de multiplicité) à l'intérieur du disque ouvert unité, ce qui montrera que cet espace vectoriel est de dimension p.

Comme les coefficients de f sont tous positifs, on a $|f(z)| \leq f(|z|)$. On a $f(1) = 1$ et $f'(1) = m > 0$, donc il existe $\epsilon > 0$, tel que pour tout $r \in [1 - \epsilon, 1[$, $f(r) \leq 1 - (1 - r)m/2 < 1$. Prenons donc donc un tel r, en demandant en plus que r dépasse le plus grand module des zéros de $1 - f$ à l'intérieur du disque unité ouvert.

Sur le cercle C_r centré en o et de rayon r, on a

$$\forall z \in C_r \quad |(1 - f(z)) - 1| = |f(z)| \leq f(r) < 1,$$

soit

$$\forall z \in C_r \quad |(1 - f(z))z^p - z^p| < |z^p|.$$

Alors, d'après le théorème de Rouché, les polynômes $Q = (1 - f(z))z^p$ et z^p ont le même nombre de zéros (comptés avec leur ordre de multiplicité) à l'intérieur de C_r, soit p zéros. Il en est de même pour l'équation $f(z) = 1$.

Pour $1 \leq i \leq p$, notons ϕ_i la forme linéaire sur E : $u \mapsto u_{-i}$. Comme E est de dimension p, pour montrer que le système $\phi_i = 1, \dots \phi_p = 1$ a une unique solution sur E, il suffit de montrer que la famille ϕ_i est libre. Cependant, E est aussi l'espace vectoriel des suites vérifiant la récurrence : $\prod_{i\in I}(X-\lambda_i)^{n_i}(\theta).u = 0$. Comme ce polynôme est de degré p, avec un terme de plus bas degré non nul (o n'est pas racine), la récurrence et la donnée de $u_{-p}, \dots, u_{-1}$ permet de déterminer u de proche en proche, ce qui montre bien que la famille des ϕ_i est libre.

On a donc bien montré que, dans tous les cas, le système d'équations était exactement déterminé et permettait donc d'identifier v.

Une copie a trouvé une solution qui évite Rouché, d'une efficacité redoutable. Je la reformule ici, en replaçant les ingrédients utilisés dans un cadre plus général.

Supposons acquis que la solution $(v_n)_{n\geq -p}$ cherchée vérifie

— v est dans l'espace vectoriel des suites $(v_n)_{n\geq -p}$ qui s'écrivent

$$v_n = \sum_{i\in I} P_i(n)\lambda_i^{n_i},$$

avec $\deg P_i \leq n_i$, qui est lui-même un sous-espace vectoriel des solutions de $(I - f(\theta))v = 0$ vérifiant $\lim v_n = 0$.

— pour tout $i \in \{1, \dots, p\}$, $v_{-i} = 1$.

Tout ceci nous donne un système linéaire, avec comme inconnues les coefficients des P_i. Comme pour tout système linéaire avec un nombre fini d'équations, dire qu'on peut le résoudre, c'est exactement dire qu'il admet une unique solution.

Pour $(w_n)_{n\geq -p}$ une telle suite, on doit montrer que w coïncide avec v. On la prolonge en posant $w_i = 1$ pour $i < -p$. Notons que la suite $(w_n)_{n\in\mathbb{Z}}$ est bornée.

Soit w une suite quelconque bornée. On note P l'opérateur associé à la chaîne de Markov (S_n), soit $Pu = f(\theta)u$. Soit $k \geq 0$.

Il est bien connu (voir par exemple le corollaire 21 page 158 de mon livre) que pour une chaîne de Markov (S_n) d'opérateur de transfert P, la suite (Y_n) définie par

$$Y_{n+1} = w(S_{n+1}) - (Pw)(S_n)$$

est une différence de martingale sous $\mathbb{P}^k$. Ainsi

$$Z_n = \sum_{i=0}^{n-1} w(S_{i+1}) - (Pw)(S_i)$$

est une martingale.

On note que

$$\forall k \in \mathbb{N} \quad (Pw)(k) = k.$$

On dit que sur $\mathbb{N}$, la fonction w est harmonique pour l'opérateur P.

Si on note T le temps d'entrée dans $]-\infty, \dots, -1]$ (c'est à dire le temps de sortie de $\mathbb{N}$), $Z_{n\wedge T}$ est encore une martingale sous $\mathbb{P}^k$ (martingale arrêtée). Sous $\mathbb{P}^k$, $S_n \wedge T$ prend presque sûrement des valeurs dans $\{-p, \dots, \}$. Ainsi $Pw(S_i) = w(S_i)$ pour $i < T$, et on en déduit que

$$Z_{n\wedge T} = w(S_{n\wedge T}) - w(S_0).$$

(somme télescopique). Ainsi, $w(S_{n\wedge T})$ est une martingale.

Sur l'événement $\{T < +\infty\}$, $w(S_{n\wedge T})$ converge vers 1, tandis que sur $\{T = +\infty\}$, $w(S_{n\wedge T})$ converge vers 0 car $S_n \to +\infty$ et $\lim w_n = 0$. Ainsi $w(S_{n\wedge T})$, converge presque sûrement vers $frm[o]--_{\{T<+\infty\}}$. Comme la suite w est bornée, par convergence dominée $\mathbb{E}^k(w(S_{n\wedge T}))$

converge vers $\mathbb{P}^k(T < +\infty) = v_k$. Or $w(S_{n\wedge T})$ est une martingale, donc son espérance est constante :

$$\mathbb{E}^k(w(S_{n\wedge T})) = \mathbb{E}^k(w(S_{0\wedge T})) = \mathbb{E}^k(w(S_0)) = w(k).$$

Cela donne le résultat voulu.

Olivier Garet, Université de Lorraine

Problème 7 : Systèmes dynamiques

Auteur : Alexei Glutsyuk (CNRS et Ecole Normale supérieure de Lyon)

On appelle *cycle limite* d'un champ de vecteurs défini sur un domaine (c'est-à-dire un ouvert connexe) de $\mathbb{R}^2$ une orbite fermée non réduite à un point, qui est isolée : cela signifie qu'elle a un voisinage dans lequel il n'y a aucune autre orbite périodique.

On se donne une fonction holomorphe $f(x+iy) = f_1(x,y)+if_2(x,y)$ sur un domaine dans $\mathbb{R}^2 = \mathbb{C}$, et on considère le champ de vecteurs $X = (f_1, f_2)$ défini sur ce domaine.

Démontrer que le champ X ne peut pas avoir de cycles limites.

Rapport

Une copie contient une solution plus directe que celle que j'avais imaginée, en montrant que les champs f et if commutent. Cela contourne la demonstration de l'holomorphie du flot. Une autre copie démontre l'holomorphie en utilisant le fait que la dérivée du flot obéït à l'équation aux variations.

Solution

Raisonnons par l'absurde. Soit le champ X a un cycle limite C. C'est donc une orbite périodique. Notons T sa période. Alors l'application du flot $g = g_v^T$ en temps T du champ X est bien définie dans un voisinage connexe V du cycle C. Le champ X étant holomorphe (c'est à dire, défini par une fonction holomorphe comme ci-dessus), l'application du flot g l'est aussi. Mais tout point de la courbe C est un point fixe pour g, par la définition de la période T. Donc, g est une application holomorphe $g : V \to \mathbb{C}$, qui fixe les points dans une courbe C. Donc, $g(z) - z \equiv 0$ le long de C. Cela implique, que $g(z) - z \equiv 0$, par le théorème d'unicité de prolongement analytique. Autrement dit, $g \equiv Id$. Donc, l'application g du flot en temps T est identité au voisinage V de C, et pour le champ de vecteurs X, l'orbite de tout point $p \in V$ est fermée et de période au plus T. Donc, toute orbite d'un point suffisamment proche de C est périodique, et C ne peut pas être un cycle limite.

Alexei Glutsyuk, CNRS et Ecole Normale supérieure de Lyon

Problème 8 : Théorie de la mesure

Auteur : Ai-Hua Fan (Université d'Amiens)

Soit $n \geq 2$ un entier, $a_1, a_2, \ldots, a_n$ des nombres réels et $b_1, b_2, \ldots, b_n$ des nombres strictement positifs. Supposons que

$$a_1 + a_2 + \cdots + a_n = 0.$$

1. Démontrer l'inégalité

$$\left| \sum_{k=1}^{n} a_k b_k \right| \leq \frac{M-m}{M+m} \sum_{k=1}^{n} |a_k b_k|, \tag{35}$$

où

$$m = \min_{1 \leq k \leq n} b_k \quad \text{et} \quad M = \max_{1 \leq k \leq n} b_k.$$

2. Montrer que la constante $\frac{M-m}{M+m}$ dans l'inégalité (35) est optimale.
3. Plus généralement, soient $(X, \mathcal{A}, \mu)$ un espace mesuré et $f \in L^1(\mu)$, $g \in L^\infty(\mu)$ deux fonctions réelles. Supposons que $\int_X f(x)\,\mathrm{d}\mu(x) = 0$, et $\alpha \leq g(x) \leq \beta$ μ-presque partout avec des constantes vérifiant $0 < \alpha < \beta < \infty$. Montrer que

$$\left| \int_X f g \,\mathrm{d}\mu \right| \leq \frac{\beta - \alpha}{\beta + \alpha} \int_X |f g| \,\mathrm{d}\mu.$$

Rapport

34 candidats ont répondu. On a privilégié la rigueur, la clarté et la simplicité, sans accorder d'importance à la concision. Il y a des copies de 7 ou 8 pages, et d'autres de 1 ou 2 pages. Certaines sont faciles à lire, d'autres plus difficiles à suivre. Certaines copies auraient mérité une relecture plus soigneuse par leurs auteurs, car les coquilles y sont nombreuses.

Une copie mérite une attention particulière, elle donne une preuve basée sur l'inégalité de Bernstein (la norme sup de la dérivée d'un polynôme trigonométrique est majorée par la norme sup du polynôme multipliée par le degré). C'est une méthode intéressante, à laquelle l'auteur n'avait pas pensé. Une note de 9 lui a été donnée, et non 10, car la preuve pour le point (3) n'est pas parfaite.

Solution

(iii) En changeant f en $-f$ si nécessaire, il suffit de montrer que

$$\int f g \,\mathrm{d}\mu \leq \frac{\beta - \alpha}{\beta + \alpha} \int |f g| \,\mathrm{d}\mu.$$

En posant

$$P := \{f > 0\} \quad \text{et} \quad N := \{f < 0\},$$

l'inégalité est équivalente à

$$(\beta + \alpha)\left(\int_P f g \,\mathrm{d}\mu + \int_N f g \,\mathrm{d}\mu \right) \leq (\beta - \alpha)\left(\int_P f g \,\mathrm{d}\mu - \int_N f g \,\mathrm{d}\mu \right),$$

ou encore à

$$2\alpha \int_P f g \, \mathrm{d}\mu + 2\beta \int_N f g \, \mathrm{d}\mu \leq 0.$$

Elle résulte de l'estimation suivante :

$$\alpha \int_P f g \, \mathrm{d}\mu + \beta \int_N f g \, \mathrm{d}\mu \leq \alpha\beta \int_P f \, \mathrm{d}\mu + \alpha\beta \int_N f \, \mathrm{d}\mu = \alpha\beta \int f \, \mathrm{d}\mu = 0.$$

Notre raisonnement montre aussi que l'égalité a lieu si et seulement si la dernière inégalité est une égalité, c'est-à-dire, si

$$g = \beta \text{ si } f > 0, \quad \text{et} \quad g = \alpha \text{ si } f < 0.$$

(i) et (ii) Appliquer (iii) à $X = \{1, 2, \ldots, n\}$ avec la mesure de comptage.

Remarque. Le résultat reste vrai si f est complexe. Soit ϕ l'argument de l'intégrale $I := \int f g \, \mathrm{d}\mu$. Alors

$$|I| = \mathrm{Re} \int e^{-i\phi} f g \, \mathrm{d}\mu = \int \mathrm{Re}(e^{-i\phi} f) g \, \mathrm{d}\mu \leq \frac{\beta - \alpha}{\beta + \alpha} \int |f g| \, \mathrm{d}\mu.$$

Ai-Hua Fan, Université d'Amiens

Problème 9 : Théorie des nombres

Auteur : Bruno Deschamps (Université du Maine)

Pour un entier $n \geq 2$ donné, déterminer l'entier minimal $\sigma(n)$ tel que tout élément de $\mathbb{Z}/n\mathbb{Z}$ soit une somme de $\sigma(n)$ carrés.

Rapport

Il y a eu beaucoup de propositions de solutions, justes du point de vue mathématique, et il a été difficile de classer les meilleures. Les critères d'excellence retenus ont été l'originalité et le caractère élémentaire des arguments. Plusieurs solutions proposées utilisaient de gros (voire très gros) théorèmes tels que les théorèmes des deux et trois carrés, le lemme de Hensel, le théorème de progression arithmétique de Dirichlet. Nous avons donc considéré que *l'utilisation de bulldozers pour écraser une mouche* était un critère discriminant pour le palmarès, car peu en accord avec la tradition de subtilité en théorie des nombres. Comme nous allons le voir, il existait plusieurs façons totalement élémentaires d'arriver au résultat, et plusieurs d'entre elles étaient concises et élégantes. Certaines de ces solutions sont même présentables à un niveau L1-L2, ce qui est remarquable.

Solutions

Pour ce qui est des preuves proposées, elles ont toutes, peu ou prou, suivi la même stratégie. Nous avons toutefois été à la fois surpris et ravis de découvrir une grande diversité d'idées avancées pour la résolution du problème. C'est sur les points clé qu'il y a eu des variantes d'arguments. Nous avons essayé dans la suite de synthétiser et présenter les différentes bonnes idées que nous

avons rencontrées. Les participants au concours devraient ainsi tous pouvoir se retrouver dans ce qui suit.

Étape 1 : si $f : A \longrightarrow B$ est un morphisme d'anneaux et si $A_0 \subset A$ est une partie où tout élément est une somme de k carrés d'éléments de A et telle que $f(A_0) = B$, alors tout élément de B est une somme de k carrés.

Une conséquence immédiate de ce résultat est que la suite $(\sigma(n))_n$ est croissante (pour l'ordre de divisibilité sur $\mathbb{N}^*$).

Étape 2 : le théorème des quatre carrés "*tout entier naturel est somme de quatre carrés*". En combinant ce résultat avec celui de l'étape 1, on en déduit que $\sigma(n) \leq 4$ pour tout $n \geq 2$. Nous verrons plus loin que l'utilisation du théorème des quatre carrés (qui est loin d'être trivial) n'était en fait pas nécessaire pour montrer que $\sigma(n) \leq 4$.

Étape 3 : le théorème des restes chinois "*si $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ désigne la décomposition en facteurs premiers de l'entier $n \geq 2$ alors les anneaux $\mathbb{Z}/n\mathbb{Z}$ et $\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_k^{\alpha_k}\mathbb{Z}$ sont isomorphes*". Une conséquence pour notre problème est alors que $\sigma(n) = \max_i \sigma(p_i^{\alpha_i})$. Le problème se ramenait donc à étudier $\sigma(p^n)$ pour tout premier p et tout entier $n \geq 1$.

Étape 4 : le cas $p = 2$. Un petit calcul montre que $\sigma(2) = 1$, $\sigma(4) = 3$ (3 n'est pas somme de deux carrés dans $\mathbb{Z}/4\mathbb{Z}$), $\sigma(8) = 4$ (7 n'est pas somme de trois carrés dans $\mathbb{Z}/8\mathbb{Z}$). Les résultats des étapes 1 et 2 montrent que l'on a ensuite $\sigma(2^n) = 4$ pour $n \geq 2$.

Étape 5 : le cas $p \neq 2$.

5.1. Cas $n = 1$. Le dénombrement classique des carrés de $\mathbb{F}_p$ montre qu'il existe des éléments de $\mathbb{F}_p$ qui ne sont pas des carrés mais que tout élément de $\mathbb{F}_p$ est une somme de deux carrés. Ainsi, $\sigma(p) = 2$ et, en vertu du résultat de l'étape 1, $\sigma(p^n) \geq 2$ pour tout $n \geq 1$.

5.2. Etude de $(\mathbb{Z}/p^n\mathbb{Z})^*$. Ce point est le cœur de la preuve. Même si cela n'a pas été toujours mentionné clairement, tout le monde a tenté de démontrer que

Tout élément de $(\mathbb{Z}/p^n\mathbb{Z})^*$ est une somme de deux carrés.

Pour démontrer cette propriété, beaucoup ont invoqué une propriété de relèvement pour affirmer que *si $a \in \mathbb{Z}$ est tel que p ne divise pas a alors $a \bmod(p)$ est un carré si est seulement si $a \bmod(p^n)$ l'est aussi*. Il y eu trois écoles :

L'école hensélienne : on considère le polynôme $P(x) = x^2 - a \in \mathbb{Z}_p[x]$ qui, une fois réduit modulo p, possède par hypothèse une racine $y \in \mathbb{Z}/p\mathbb{Z}$. Puisque $P'(y) = 2y \neq 0$, le lemme d'Hensel permet d'affirmer qu'il existe $\alpha \in \mathbb{Z}_p$ tel que $P(\alpha) = 0$. Si pour $n \geq 2$, on note α_n l'image de α par l'épimorphisme $\mathbb{Z}_p \longrightarrow \mathbb{Z}/p^n\mathbb{Z}$, alors $\alpha_n^2 = a \bmod(p^n)$.

———

L'école élémentaire : on opère une récurrence sur $n \geq 1$. Supposons qu'il existe $y \in \mathbb{Z}$ tel que $y^2 \equiv a \bmod(p^n)$. Il existe donc $k \in \mathbb{Z}$ tel que $a = y^2 + kp^n$ et puisque p ne divise visiblement pas y (par hypothèse p ne divise pas a) l'élément $y \bmod(p^{n+1})$ est inversible dans $\mathbb{Z}/p^{n+1}\mathbb{Z}$. Puisque $p \neq 2$, $2y \bmod(p^{n+1})$ est aussi inversible. On a alors

$$\left(y \bmod(p^{n+1}) + \left(2y \bmod(p^{n+1})\right)^{-1} \left(kp^n \bmod(p^{n+1})\right)\right)^2 = a \bmod(p^{n+1}) \tag{36}$$

———

L'école combinatoire : il s'agit de montrer que, pour $n \geq 2$,

$$(\mathbb{Z}/p^n\mathbb{Z})^{*2} = \left\{a \bmod(p^n)/\ a \bmod(p) \in (\mathbb{Z}/p\mathbb{Z})^{*2}\right\} \tag{37}$$

(ici le carré d'ensemble désigne l'ensemble des carrés et non le produit cartésien). L'inclusion directe est claire. Puisque les fibres de l'épimorphisme $\mathbb{Z}/p^n\mathbb{Z} \longrightarrow \mathbb{Z}/p\mathbb{Z}$ compte exactement p^{n-1} éléments et que $(\mathbb{Z}/p\mathbb{Z})^{*2}$ en compte $\frac{p-1}{2}$, il vient que

$$\sharp\left\{a \bmod(p^n)/\ a \bmod(p) \in (\mathbb{Z}/p\mathbb{Z})^{*2}\right\} = \frac{p^{n-1}(p-1)}{2} \tag{38}$$

On va voir un peu plus loin que $\sharp(\mathbb{Z}/p^n\mathbb{Z})^{*2} = \frac{p^{n-1}(p-1)}{2}$. L'égalité annoncée en découle alors.

———

Scholie. *Les arithméticiens des corps en herbe verront que cette propriété de relèvement permet de montrer que le niveau du corps $\mathbb{Q}_p$ (le nombre minimal $\nu(\mathbb{Q}_p)$ tel que -1 soit une somme de $\nu(\mathbb{Q}_p)$ carrés dans $\mathbb{Q}_p$) vaut 1 ou 2.*

Une fois établie cette propriété de relèvement, on considère un entier x tel que $p \nmid x$. Il existe $a, b \in \mathbb{Z}$ tels que $x \equiv a^2 + b^2 \bmod(p)$. Si $b \equiv 0 \bmod(p)$ alors x est un carré non nul modulo p et il l'est donc modulo p^n en vertu de ce qui précède. Si $b \not\equiv 0 \bmod(p)$ alors $x - a^2$ est un carré non nul modulo p et il l'est donc modulo p^n. Ainsi, il existe $c \in \mathbb{Z}$ tel que $x \equiv a^2 + c^2 \bmod(p^n)$. Ceci prouve finalement que $x \bmod(p^n)$ est bien une somme de deux carrés.

———

Il était possible de montrer que tout inversible modulo p^n est une somme de deux carrés sans utiliser un lemme de relèvement. Dans sa solution initiale, l'auteur du sujet avait proposé une preuve combinatoire directe, preuve dont les idées ne sont pas d'intersection vide avec celles de l'école combinatoire citée plus haut.

Une preuve combinatoire directe : On note respectivement $\boxed{1}$ et $\boxed{2}$ les carrés et les sommes de deux carrés d'éléments de $\mathbb{Z}/p^n\mathbb{Z}$.

Commençons par dénombrer les carrés dans $(\mathbb{Z}/p^n\mathbb{Z})^*$. L'application $c : x \longmapsto x^2$ est un endomorphisme du groupe $(\mathbb{Z}/p^n\mathbb{Z})^*$ et son image est bien évidemment égale à $(\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{1}$.

Puisque $(\mathbb{Z}/p^n\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/p^{n-1}\mathbb{Z}$, on voit que $(\mathbb{Z}/p^n\mathbb{Z})^*$ possède un unique élément d'ordre 2 et donc, le noyau $\ker(c)$ est composé de seulement deux éléments. Le premier théorème d'isomorphisme nous permet alors d'affirmer que

$$\sharp(\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{1} = \frac{(p-1)p^{n-1}}{2} \tag{39}$$

Considérons maintenant l'ensemble $G = (\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{2}$. L'ensemble $\boxed{2}$ est un monoïde multiplicatif, mais puisque, si $x^2 + y^2$ est inversible, alors

$$\frac{1}{x^2+y^2} = \frac{x^2}{(x^2+y^2)^2} + \frac{y^2}{(x^2+y^2)^2} \in \boxed{2} \tag{40}$$

on en déduit que G est un sous-groupe de $(\mathbb{Z}/p^n\mathbb{Z})^*$. Il est clair que $(\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{1} \subset (\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{2}$ et donc l'on a donc $o(G) \geq \frac{(p-1)p^{n-1}}{2}$. Il vient alors

$$[(\mathbb{Z}/p^n\mathbb{Z})^* : G] \leq 2 \tag{41}$$

Si $[(\mathbb{Z}/p^n\mathbb{Z})^* : G] = 2$, on a alors $G = (\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{1}$ et, puisque $1^2 + 1^2 = 2 \in \boxed{1}$, on en déduit par récurrence que, pour tout $i = 1, \cdots, p-1$, on a $i = 1 + (i-1) \in \boxed{1}$. Maintenant, $p+1 = (p-1)+2$ et

donc $(p+1) \in \boxed{1}$. Par récurrence, on montre ainsi que pour tout k premier à p, $k \in \boxed{1}$, c'est-à-dire que $(\mathbb{Z}/p^n\mathbb{Z})^* \subset \boxed{1}$. Ceci est absurde d'après ce qui précède.

Ainsi, $[(\mathbb{Z}/p^n\mathbb{Z})^* : G] = 1$ et donc $(\mathbb{Z}/p^n\mathbb{Z})^* = (\mathbb{Z}/p^n\mathbb{Z})^* \cap \boxed{2}$, ce qui prouve finalement que $(\mathbb{Z}/p^n\mathbb{Z})^* \subset \boxed{2}$.

———

Remarque : On peut appliquer l'idée précédente pour le cas $p = 2$ et montrer que $\sigma(2^n) \leq 4$ sans utiliser le théorème des quatre carrés. Notons $\boxed{4}$ l'ensemble des sommes de quatre carrés d'éléments de $\mathbb{Z}/2^n\mathbb{Z}$. L'identité quaternionique de Hamilton montre que cet ensemble est stable par produit si bien que, en appliquant le même argument que précédemment, l'ensemble $(\mathbb{Z}/2^n\mathbb{Z})^* \cap \boxed{4}$ est un sous-groupe de $(\mathbb{Z}/2^n\mathbb{Z})^*$ et l'on a

$$(\mathbb{Z}/2^n\mathbb{Z}) \cap \boxed{1} \subsetneq (\mathbb{Z}/2^n\mathbb{Z})^* \cap \boxed{2} \subsetneq (\mathbb{Z}/2^n\mathbb{Z})^* \cap \boxed{4} \subset (\mathbb{Z}/2^n\mathbb{Z})^* \tag{42}$$

Pour voir que les deux premières inclusions sont strictes, il suffit de remarquer que pour $n = 3$, l'élément 5 est un inversible qui est somme de deux carrés sans être un carré et 7 est un inversible qui est somme de quatre carrés sans être une somme de deux carrés.

On considère à nouveau l'épimorphisme $c : x \longmapsto x^2$ de $(\mathbb{Z}/2^n\mathbb{Z})^*$ sur $(\mathbb{Z}/2^n\mathbb{Z}) \cap \boxed{1}$. Puisque $(\mathbb{Z}/2^n\mathbb{Z})^* \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z}$, on voit que ce groupe compte exactement trois éléments d'ordre 2 et donc $\mathrm{Im}(c) = (\mathbb{Z}/2^n\mathbb{Z}) \cap \boxed{1}$ est d'indice $o(\ker(c)) = 4$ dans $(\mathbb{Z}/2^n\mathbb{Z})^*$. Les inclusions précédentes assure alors que $(\mathbb{Z}/2^n\mathbb{Z})^* \cap \boxed{4}$ est d'indice 1 dans $(\mathbb{Z}/2^n\mathbb{Z})^*$, c'est-à-dire que $(\mathbb{Z}/2^n\mathbb{Z})^* \cap \boxed{4} = (\mathbb{Z}/2^n\mathbb{Z})^*$.

Pour finir de montrer que $\sigma(2^n) \leq 4$, on voit qu'une puissance de 2 dans $\mathbb{Z}/2^n\mathbb{Z}$ est soit un carré, soit une somme de deux carrés et donc dans tous les cas une somme de quatre carrés. L'identité quaternionique permet alors d'en déduire que $\sigma(2^n) \leq 4$.

———

D'autres preuves directes ont été proposées :

Une autre preuve combinatoire directe : Comme vu précédemment, $(\mathbb{Z}/p^n\mathbb{Z})^{*2}$ est un sous-groupe d'indice 2 de $(\mathbb{Z}/p^n\mathbb{Z})^*$. Ainsi, en regardant les images des éléments de $(\mathbb{Z}/p^n\mathbb{Z})^*$ dans le groupe quotient $(\mathbb{Z}/p^n\mathbb{Z})^*/(\mathbb{Z}/p^n\mathbb{Z})^{*2}$, on en déduit que dans $(\mathbb{Z}/p^n\mathbb{Z})^*$, le produit de deux éléments dont aucun n'est un carré, est un carré.

Une fois cette remarque faite, on raisonne par l'absurde en supposant qu'il existe un élément $x \in (\mathbb{Z}/p^n\mathbb{Z})^*$ qui ne soit pas somme de deux carrés (x n'est en particulier pas un carré). On a donc, pour tout $a, b \in \mathbb{Z}/p^n\mathbb{Z}$, $x \neq a^2(1+b^2) = a^2 + (ab)^2$. Ainsi,

$$\forall b \in \mathbb{Z}/p^n\mathbb{Z},\ 1 + b^2 \in (\mathbb{Z}/p^n\mathbb{Z})^* \Longrightarrow 1 + b^2 \in (\mathbb{Z}/p^n\mathbb{Z})^{*2} \tag{43}$$

car si $1 + b^2 \in (\mathbb{Z}/p^n\mathbb{Z})^* - (\mathbb{Z}/p^n\mathbb{Z})^{*2}$ alors $x(1+b^2)^{-1} \in (\mathbb{Z}/p^n\mathbb{Z})^{*2}$ d'après la remarque précédente.

Puisque $2 = 1 + 1^2$, 2 est un carré et, par récurrence, on en déduit que tout $r = 1, \cdots, p-1$ est un carré. Par ailleurs, puisque 2 est un carré, pour tout $a, b \in \mathbb{Z}/p^n\mathbb{Z}$, on a $x \neq a^2(2+b^2) = (\sqrt{2}a)^2 + (ab)^2$. Comme précédemment, on en déduit que

$$\forall b \in \mathbb{Z}/p^n\mathbb{Z},\ 2 + b^2 \in (\mathbb{Z}/p^n\mathbb{Z})^* \Longrightarrow 2 + b^2 \in (\mathbb{Z}/p^n\mathbb{Z})^{*2} \tag{44}$$

Comme $p + 1 = 2 + (p - 1)$, on en déduit que $p + 1$ est un carré et, par récurrence, que tout $r = p + 1, \cdots, 2p - 1$ est un carré. En appliquant p^{n-1}-fois ce raisonnement, on en déduit finalement que $(\mathbb{Z}/p^n\mathbb{Z})^* = (\mathbb{Z}/p^n\mathbb{Z})^{*2}$, ce qui est absurde.

———

Remarques : 1/ On peut appliquer l'idée développée ici pour donner une nouvelle preuve du fait que les éléments de $(\mathbb{Z}/2^n\mathbb{Z})^*$ sont des sommes de quatre carrés.

Les inclusions strictes établies dans le paragraphe précédent montrent que $(\mathbb{Z}/2^n\mathbb{Z})^* \cap \boxed{2}$ est un sous-groupe d'indice 2 de $(\mathbb{Z}/2^n\mathbb{Z})^*$ et l'on en déduit la propriété suivante : si $x, y \in (\mathbb{Z}/2^n\mathbb{Z})^*$ désignent deux éléments dont aucun n'est une somme de deux carrés, alors le produit xy est une somme de deux carrés.

On raisonne maintenant par l'absurde en supposant qu'il existe $x \in (\mathbb{Z}/2^n\mathbb{Z})^*$ qui ne puisse s'écrire comme somme de quatre carrés. Pour tout $a, b, c, d \in \mathbb{Z}/2^n\mathbb{Z}$, on a donc $x \neq (a^2+b^2)(1+(c^2+d^2)) = a^2+b^2+(a^2+b^2)(c^2+d^2)$ (somme de quatre carrés d'après l'identité de Jacobi). On en déduit que pour tout $c, d \in \mathbb{Z}/2^n\mathbb{Z}$, si $1+(c^2+d^2)$ est inversible alors il est somme de deux carrés. Pour $c = d = 1$ on en déduit que 3 est somme de deux carrés dans $\mathbb{Z}/2^n\mathbb{Z}$, et donc dans $\mathbb{Z}/4\mathbb{Z}$, ce qui est absurde.

———

2/ Nous avons rencontré une autre preuve, totalement élémentaire, pour la majoration $\sigma(2^n) \leq 4$. Elle repose sur la propriété suivante : *si $a \in \mathbb{Z}$ est tel que $a \equiv 1 \bmod(8)$, alors a est un carré modulo 2^n pour tout $n \geq 3$*. Pour $a = 1 + 8k$, la preuve se fait par récurrence. Pour $n = 3$ la proposition est claire. Si la propriété est vraie pour $n \geq 3$ alors, il existe $b, l \in \mathbb{Z}$ tel que $b^2 = a + l2^n$ et, pour un entier h quelconque, on a alors $(b + h2^{n-1})^2 \equiv a + (l + bh)2^n \bmod(2^{n+1})$. L'entier b étant impair, on peut choisir l'entier $h = 0, 1$ pour que $2|(l + bh)$ et donc a est bien un carré modulo 2^{n+1}.

Une fois établi ce résultat, supposons par l'absurde qu'il existe un plus petit entier $a > 0$ tel que $a \bmod(2^n)$ ne soit pas une somme de quatre carrés. On a nécessairement $a \not\equiv 0 \bmod(4)$, car sinon $a/4 < a$ est une somme de quatre carrés et donc a est somme de quatre carrés. Donc $a \equiv 1, 2, 3, 5, 6, 7 \bmod(8)$ et donc il existe un entier $b \equiv 1 \bmod(8)$ tel que $a = b, b+1, b+2, b+4, b+5, b+6$. L'élément $b \bmod(2^n)$ est un carré, mais les éléments $1, 2, 3, 5, 6, 7 \bmod(2^n)$ étant tous des sommes de trois carrés, on en déduit une contradiction.

———

Une preuve dirichletienne : Soit $a \in \mathbb{Z}$ premier à p. Puisque 4 et p sont premier entre eux, le théorème des restes chinois assure qu'il existe $b \in \mathbb{Z}$ tel que $b \equiv 1 \bmod(4)$ et $b \equiv a \bmod(p^n)$. Puisque b est premier avec $4p^n$, le théorème de progression arithmétique de Dirichlet assure qu'il existe un premier q tel que $q \equiv b \bmod(4p^n)$. Puisque $q \equiv 1 \bmod(4)$ il est somme de deux carrés (théorème des deux carrés) et comme $q \equiv a \bmod(p^n)$, a est bien somme de deux carrés modulo p^n.

———

Une preuve hensélienne : C'est exactement la même idée que pour relever les carrés. Si $a \in \mathbb{Z}$ est premier à p alors il existe $c, d \in \mathbb{Z}$ tel que $a \equiv b^2 + c^2 \bmod(p)$. Les entiers b et c ne peuvent être simultanément divisibles par p, donc par exemple, $b \not\equiv 0 \bmod(p)$. On considère alors le polynôme $P(x) = x^2 + (b^2 - a)$ et l'on applique le lemme de Hensel comme précédemment.

Certains membres de l'école hensélienne ont même utilisé une variante à plusieurs variables du lemme de Hensel, en considérant le polynôme $P(x, y) = x^2 + y^2 - a$.

———

Une fois établi le fait que tout inversible modulo p^n est une somme de deux carrés, on constate que si $x \in \mathbb{Z}$ est tel que $p \nmid x$ alors $x \bmod(p^n)$ est une somme de deux carrés et si $p|x$ alors $p \nmid (x-1)$ et donc $(x-1) \bmod(p^n)$ est une somme de deux carrés, donc $x \bmod(p^n)$ est une somme de trois carrés. En conclusion, on a $2 \leq \sigma(p^n) \leq 3$.

Remarque : Il est donc possible de n'utiliser que des propriétés élémentaires pour montrer que $\sigma(p^n) \leq 4$ pour tout premier p et donc que $\sigma(n) \leq 4$ pour tout $n \geq 2$.

5.3. Détermination de $\sigma(p^n) = 2, 3$ pour $n \geq 2$. La fin de la preuve consistait à montrer que la valeur de $\sigma(p^n)$ est entièrement déterminée par la congruence modulo 4 du premier p.

5.3.1. Cas $p \equiv 1 \bmod(4)$. Dans cette situation il fallait remarquer que $p \bmod(p^n)$ était une somme de deux carrés. Il y a eu deux écoles pour démontrer ce résultat :

L'école du théorème des deux carrés, qui invoque le fait qu'un premier p vérifiant $p \equiv 1 \bmod(4)$ est somme de deux carrés dans $\mathbb{Z}$, donc dans $\mathbb{Z}/p^n\mathbb{Z}$.

L'école élémentaire : si $p \equiv 1 \bmod(4)$ alors $(-1)^{(p-1)/2} \equiv 1 \bmod(4)$ et donc -1 est un carré modulo p (propriété classique et élémentaire du symbôle de Legendre). Il existe donc $a \in \mathbb{Z}$ tel que $p \equiv 1 + a^2 \bmod(p)$ et l'on écrit $p = 1 + a^2 + \lambda p$. Comme $p \neq 2$, 2 est inversible modulo p^2 et il existe $d \in \mathbb{Z}$ tel que $2d \equiv 1 \bmod(p^2)$. On a alors

$$(1 + d\lambda p)^2 + a^2 \equiv 1 + \lambda p + a^2 \equiv p \bmod(p^2) \tag{45}$$

et donc $p \bmod(p^2)$ est une somme de deux carrés dont l'un au moins est inversible. En appliquant alors la propriété de relèvement détaillée dans la partie 5.2. on en déduit que p est une somme de deux carrés modulo p^n.

Une fois établi ce résultat il convenait de se rappeler que, eu égard à l'identité de Jacobi

$$(a_1^2 + a_2^2)(b_1^2 + b_2^2) = (a_1 b_1 + a_2 b_2)^2 + (a_1 b_2 - a_2 b_1)^2 \tag{46}$$

les sommes de deux carrés dans un anneau sont stables par produit. Tout élément de $\mathbb{Z}/p^n\mathbb{Z}$ est le produit d'un élément de $(\mathbb{Z}/p^n\mathbb{Z})^*$ et d'une puissance de p modulo p^n. Il est donc somme de deux carrés vu ce qui précède.

On constate réciproquement que, si $p \bmod(p^n) = (a^2 + b^2) \bmod(p^n)$ est une somme de deux carrés, alors $a^2 + b^2 \equiv 0 \bmod(p)$. L'un des deux entiers a ou b n'est pas divisible par p, sinon les deux le seraient et, pour $n = 2$, on aurait $p \equiv a^2 + b^2 \equiv 0 \bmod(p^2)$, ce qui est absurde. On a donc, par exemple, $a \not\equiv 0 \bmod(p)$ (et a est alors inversible modulo p^2). On a alors

$$-1 \bmod(p) = \Big((b \bmod(p))\,(a \bmod(p))^{-1}\Big)^2 \tag{47}$$

ce qui implique que $p \equiv 1 \bmod(4)$ (réciproque de la propriété du symbôle de Legendre rappelée plus haut).

En conclusion, on a $\sigma(p^n) = 2 \Longleftrightarrow p \equiv 1 \bmod(4)$.

5.3.2 Cas $p \equiv -1 \bmod(4)$**.** Si $p \not\equiv 1 \bmod(4)$ alors comme $p \neq 2$ on a $p \equiv -1 \bmod(4)$. Ce qui précède prouve que $\sigma(p^n) = 3 \Longleftrightarrow p \equiv -1 \bmod(4)$.

Étape 6 : Conclusion. On déduit de ce qui précède que

Théorème.— *Si* $n = 2^h p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ *désigne la décomposition en facteurs premiers de l'entier* $n \geq 2$ *alors on a* $\sigma(n) \leq 4$ *et*

$\sigma(n) = 1$ *si et seulement si* $n = 2$.

$\sigma(n) = 2$ *si et seulement si* $h \leq 1$ *et, pour tout* $i = 1, \cdots, k$, $p_i \equiv -1\ (4) \Longrightarrow \alpha_i = 1$.

$\sigma(n) = 3$ *si et seulement si [*$h = 2$*] ou [(*$h \leq 1$*) et (il existe* $i = 1, \cdots, k$ *tel que* $p_i \equiv -1\ (4)$ *et* $\alpha_i \geq 2$*)].*

$\sigma(n) = 4$ *si et seulement si* $h \geq 3$.

Quelques valeurs :

n	$\sigma(n)$	n	$\sigma(n)$	n	$\sigma(n)$	n	$\sigma(n)$
2	1	12	3	22	2	32	4
3	2	13	2	23	2	33	2
4	3	14	2	24	4	34	2
5	2	15	2	25	2	35	2
6	2	16	4	26	2	36	3
7	2	17	2	27	3	37	2
8	4	18	3	28	3	38	2
9	3	19	2	29	2	39	2
10	2	20	3	30	2	40	4
11	2	21	2	31	2	41	2

Bruno Deschamps, Université du Maine

Problème 10 : Topologie

Auteur : Marc Peigné (Université de Tours)

On dit qu'un espace métrique (X,d) est *géodésique* si pour tous points x et $y \in X$, il existe (au moins) un *segment géodésique* joignant x et y, c'est-à-dire une application $\gamma : [0, d(x,y)] \to X$ telle que $\gamma(0) = x$, $\gamma(d(x,y)) = y$ et $d(\gamma(s), \gamma(t)) = |s-t|$ pour tous $0 \leq s, t \leq d(x,y)$.

Soit (X,d) un espace métrique localement compact. Soit G un groupe agissant à gauche sur X par isométries. On note simplement l'action $(g,x) \mapsto gx$. On dit que l'action est *proprement discontinue* si, pour toute partie compacte $K \subset X$, l'ensemble

$$\{g \in G\,;\, gK \cap K \neq \emptyset\}$$

est fini.

On fixe un espace métrique (X,d) géodésique localement compact. On se donne une action à gauche proprement discontinue d'un groupe G par isométries sur (X,d). On suppose que l'espace quotient X/G est compact.

1. Montrer que pour tout $x \in X$ et tout $R > 0$, le nombre

$$N_G(x,R) := \#\{g \in G\,;\, d(x, g(x)) \leq R\}$$

est fini.

2. Montrer que la limite

$$\delta_G := \lim_{R\to\infty} \frac{1}{R} \ln N_G(x,R)$$

existe, et qu'il existe une constante C telle que, pour tout $R > 0$,

$$N_G(x,R) \geq C e^{\delta_G R}.$$

Indication. Soit Λ un réel suffisamment grand pour que les boules $B(gx,\Lambda)$ recouvrent X lorsque g décrit G. On montrera que la fonction $R \mapsto \#\{g \in G\,;\, R - 2\Lambda \leq d(x, g(x)) \leq R + 2\Lambda\}$ est sous-multiplicative.

Rapport

Ce sujet est inspiré de résultats concernant les groupes d'isométries de variétés à courbure négative. L'exposant δ_G est appelé aussi exposant de Poincaré, c'est l'exposant critique de la série du même nom (critique au sens où cette série converge pour $s > \delta_G$ et diverge pour $s < \delta_G$). Une question naturelle est de savoir si cette série diverge en $s = \delta_G$, auquel cas le groupe est dit divergent. Cette question semble au premier abord anodine, elle est en fait très profonde et a des conséquences importantes sur la dynamique du flot géodésique. Pour les groupes co-compacts, l'argument de sous-additivité, objet du présent énoncé, est la façon la plus simple de répondre positivement à cette question.

Une copie a développé une approche personnelle de la question 2, en passant par la métrique des mots. Ce passage n'est pas naturel dans le contexte de l'énoncé, même si pour la métrique des mots la réponse est plus rapide ; mais je dois avouer que j'ai été surpris qu'on puisse revenir à la distance initiale en introduisant un paramètre supplémentaire, cela m'a beaucoup intrigué et intéressé.

Solution

1. Soit $x \in X$ et $R > 0$. Pour pouvoir appliquer l'hypothèse d'action proprement discontinue, il suffirait de montrer que $\bar{B}(x, R)$ est compacte. On commence par cela.

Soit $J \subset \mathbb{R}_+$ l'ensemble des rayons R tels que $\bar{B}(x, R)$ est compacte. C'est un intervalle contenant 0.

Montrons que J est ouvert. Soit $R \in J$. Comme X est localement compact, pour tout $y \in X$, il existe un rayon r_y tel que $\bar{B}(y, r_y)$ soit compacte. Les boules ouvertes $B(y, \frac{1}{2} r_y)$ constituent un recouvrement ouvert de $\bar{B}(x, R)$, qui est compacte par hypothèse. Extrayons un recouvrement fini $B(y_j, \frac{1}{2} r_j)$. Soit δ le plus petit des rayons $\frac{1}{2} r_j$. Soit $z \in \bar{B}(x, R + \delta)$. Sur un segment géodésique γ reliant x à z, on trouve le point $z' = \gamma(R) \in \bar{B}(x, R)$ qui satisfait $d(z', z) \leq \delta$. Par construction, il existe j tel que $d(y_j, z') < \frac{1}{2} r_j$. Alors $d(z, y_j) < \frac{1}{2} r_j + \delta < r_j$. Cela prouve que $\bar{B}(x, R + \delta) \subset \bigcup_j \bar{B}(y_j, r_j)$. Cette réunion finie de compacts est compacte, $\bar{B}(x, R + \delta)$ est fermée donc $\bar{B}(x, R + \delta)$ est compacte. Cela démontre que $[0, R + \delta[\subset J$, J est un voisinage de R.

Montrons que J est fermé. Soit R_m sa borne supérieure. Si $R_m = +\infty$, $J = \mathbb{R}_+$ est fermé. Supposons que R_m est fini. Considérons l'ensemble Γ des segments géodésiques $\gamma : [0, R_m] \to X$ issus de x. Pour la topologie de la convergence simple sur $[0, R_m[$, c'est un compact car, par construction, pour tout $t < R_m$, $\gamma(t) \in \bar{B}(x, t)$ qui est compacte, et la condition d'isométrie passe à la limite simple. Comme les éléments de Γ sont équicontinus sur $[0, R_m]$, la convergence simple sur $[0, R_m[$ entraîne la convergence uniforme sur $[0, R_m]$. L'application $\Gamma \to X$, $\gamma \mapsto \gamma(R_m)$ est donc continue. Son image est un compact de X. Par hypothèse, il contient $\bar{B}(x, R_m)$, qui est fermée, donc $\bar{B}(x, R_m)$ est compacte. Cela démontre que $R_m \in J$, donc J est fermé.

Par connexité de $\mathbb{R}_+$, on conclut que $J = \mathbb{R}_+$, toutes les boules fermées de X sont compactes.

Si $g \in G$ et $d(x, gx) \leq R$, alors $\bar{B}(x, R) \cap g\bar{B}(x, R) \neq \emptyset$. Or $\bar{B}(x, R)$ est compacte, l'ensemble des éléments g de G tels que $\bar{B}(x, R) \cap g\bar{B}(x, R) \neq \emptyset$ est fini, donc $\{g \in G\,;\, d(x, gx) \leq R\}$ est fini.

2. La fonction

$$y \mapsto \inf\{d(gx, y)\,;\, g \in G\}$$

est continue sur X, car elle est 1-lipschitzienne. Elle passe au quotient en une fonction f continue sur l'espace des orbites X/G. Par hypothèse, X/G est compact, donc f est bornée. Choisissons un réel $\Lambda > 2 \sup f$. Soit $y \in X$. Comme $\Lambda > \sup f$, il existe $g \in G$ tel que $d(gx, y) \leq \Lambda$. Alors $y \in \bar{B}(x, \Lambda)$. Cela démontre que les images de $\bar{B}(x, \Lambda)$ par les éléments de G recouvrent X.

Pour n entier naturel, notons

$$E_n = \{g \in G\,;\, (n-2)\Lambda \leq d(x, gx) \leq (n+2)\Lambda\}$$

et $u_n = \#E_n$. On va montrer que la suite (u_n) est sous-multiplicative, i.e. pour tous entiers m et n, $u_{m+n} \leq u_m u_n$.

Soit $k \in E_{m+n}$ avec $m+n > 0$. Posons $d(x,kx) = (n+m)\Lambda + 2\lambda$ with $-\Lambda \leq \lambda \leq \Lambda$. Soit γ un segment géodésique reliant x à kx. Posons $y = \gamma(m\Lambda + \lambda)$ et notons que $d(x,y) = m\Lambda + \lambda$ et $d(y,kx) = n\Lambda + \lambda$.

Il existe $g \in G$ tel que $d(y,gx) \leq \Lambda$. Alors

$$d(x,gx) \geq d(x,y) - d(y,gx) \geq m\Lambda + \lambda - \Lambda \geq (m-2)\Lambda$$

et

$$d(x,gx) \leq d(x,y) + d(y,gx) \leq m\Lambda + \lambda + \Lambda \leq (m+2)\Lambda$$

Posons $h = g^{-1}k$. Alors $d(x,hx) = d(x,g^{-1}kx) = d(gx,kx)$, car g est une isométrie. Il vient

$$d(x,hx) \leq d(gx,y) + d(y,kx) \leq \Lambda + n\Lambda + \lambda \leq (n+2)\Lambda$$

et

$$d(x,hx) \geq -d(gx,y) + d(y,kx) \geq -\Lambda + n\Lambda + \lambda \geq (n-2)\Lambda.$$

Le procédé (qui fait intervenir des choix, mais ça n'a pas d'importance) fournit une application $E_{m+n} \to E_m \times E_n$, $k \mapsto (g,h)$. Elle est injective, puisque $k = gh$. Par conséquent

$$u_{m+n} \leq u_m u_n.$$

Posons $v_n = \ln u_n$, on a $v_n \geq 0$ car $u_n \geq 1$. La suite (v_n) est sous-additive,

$$v_{m+n} \leq v_m + v_n.$$

Il en résulte que la limite

$$L := \lim_{n\to\infty} \frac{1}{n} v_n = \inf_{n\geq 1} \frac{1}{n} v_n$$

existe et est ≥ 0. En particulier, pour tout $n \geq 1$, $u_n \geq e^{nL}$.

Pour $n \geq 2$, on a $\{g \in G\,;\, d(x,gx) \leq n\Lambda\} \supset E_{n-2}$, d'où

$$\#\{g \in G\,;\, d(x,gx) \leq n\Lambda\} \geq u_{n-2} \geq e^{(n-2)L} = C\, e^{nL},$$

où $C = e^{-2L}$.

Inversement, fixons $L' > L$; il existe M tel que

$$n \geq M \Rightarrow u_n < e^{nL'}.$$

Pour tout $n \geq M$, l'inclusion

$$\{g \in G\,;\, d(x,gx) \leq n\Lambda\} \subset \bigcup_{m\leq n-1} E_m$$

entraîne

$$\#\{g \in G\,;\, d(x,gx) \leq n\Lambda\} \leq \sum_{m\leq n-1} u_m \leq \frac{e^{nL'}}{e^{L'}-1} + D \leq C'\, e^{nL'} + D,$$

où $D = \sum_{k=1}^{M-1} u_k$ et $C' = \frac{1}{e^{L'}-1}$ ne dépendent pas de n.

Si $R > 0$ est un réel quelconque, soit $n = \lfloor \frac{R}{\Lambda} \rfloor$ la partie entière de $\frac{R}{\Lambda}$. Par construction,

$$n\Lambda \leq R < (n+1)\Lambda.$$

Donc

$$\begin{aligned} N_G(x,R) &= \#\{g\in G\,;\, d(x,gx)\le R\} \\ &\ge \#\{g\in G\,;\, d(x,gx)\le n\Lambda\} \\ &\ge C\,e^{nL} \\ &\ge C\,e^{(\frac{R}{\Lambda}-1)L}. \end{aligned}$$

Posons $\delta_G := \frac{L}{\Lambda}$. Il vient

$$N_G(x,R)\ge C''\,e^{R\delta_G},$$

où $C''=C\,e^{-L}$. Inversement, fixons $\delta'>\delta_G$. Alors $L'=\Lambda\delta'>L$.

$$N_G(x,R)\le \#\{g\in G\,;\, d(x,gx)\le (n+1)\Lambda\}\le C'\,e^{(n+1)L'}+D\le C'''\,e^{R\delta'}+D,$$

où $C'''=C'e^{L}$, D dépendent de δ' et non de R. Cet encadrement entraine que

$$\lim_{R\to\infty}\frac{1}{R}\ln N_G(x,R)=\delta_G.$$

Marc Peigné, Université de Tours