CAHIERS DU SÉMINAIRE D'HISTOIRE DES MATHÉMATIQUES

JEAN-PIERRE KAHANE

Le 13ème problème de Hilbert : un carrefour de l'algèbre, de l'analyse et de la géométrie

Cahiers du séminaire d'histoire des mathématiques 1re série, tome 3 (1982), p. 1-25

<http://www.numdam.org/item?id=CSHM_1982__3__1_0>

LE 13ème PROBLEME DE HILBERT : un carrefour de l'algèbre, de l'analyse et de la géométrie*

par Jean-Pierre Kahane

Les lecteurs de Normat connaissent l'importance des 23 problèmes posés par David Hilbert, en 1900, au Congrès International des Mathématiciens à Paris[1]. Je voudrais les entretenir du 13ème problème et de la solution de Kolmogorov.

L'origine du problème est la théorie des équations algébriques. Jusqu'au degré 6 les équations algébriques peuvent être résolues au moyen de fonctions de deux variables. L'équation générale de degré 7 peut être réduite à

$$X^7 + xX^3 + yX^2 + zX + 1 = 0 \tag{1}$$

au moyen d'additions et de fonctions d'une variable. La solution de l'équation (1) est une fonction des 3 variables x, y, z - comme l'équation (1) admet 7 racines, c'est d'ailleurs une fonction multiforme -. Hilbert demande de montrer que l'équation (1) ne peut pas être résolue au moyen de fonctions continues de deux variables seulement par une chaine finie de substitutions. Et il ajoute qu'il sait établir l'existence de fonctions analytiques de trois variables x, y, z qui ne peuvent pas s'obtenir au moyen de fonctions de deux variables seulement.

Se fiant à cet énoncé de Hilbert, les mathématiciens ont généralement cru, pendant 50 ans, qu'il existait de bonnes fonctions de trois variables réelles x, y, z,

* Article destiné à une traduction en suédois pour Normat.

qui ne pouvaient pas s'écrire en composant des fonctions continues de deux variables réelles. Certains ont même cru l'avoir démontré. La surprise a donc été grande, en 1957, lorsque le jeune mathématicien soviétique V. I. Arnold publia que toute fonction continue de trois variables peut s'écrire sous la forme

$$(2) \qquad f(x,y,z) = \sum_{j=1}^{9} f_j(\varphi_j(x,y),z)$$

où les f_j et les φ_j sont des fonctions continues.[2] Quelques semaines plus tard, un mathématicien soviétique très célèbre, A. N. Kolmogorov, qui avait déjà beaucoup réfléchi au problème (Arnold utilise certaines de ses idées) publia un résultat encore plus remarquable : toute fonction continue de n variables, définie sur le cube unité de R^n, peut s'écrire

$$(3) \qquad f(x_1, x_2, \dots x_n) = \sum_{q=1}^{2n+1} g_q\Big(\sum_{p=1}^{n} \varphi_{pq}(x_p)\Big),$$

où les fonctions φ_{pq} sont monotones, continues, et indépendantes de f, et les fonctions g_q continues (et dépendant de f, naturellement).[3] Ainsi toute fonction continue de n variables s'obtient au moyen de fonctions d'une variable, et d'additions, par une chaîne finie de substitutions.

Le problème de Hilbert se trouve donc résolu par la négative.

Du même coup, d'autres problèmes émergent :

1. Que voulait dire Hilbert en affirmant l'existence de fonctions analytiques de trois variables non représentables au moyen de fonctions de deux variables ? Ici la réponse est claire ; il s'agissait certainement de fonctions analytiques de deux variables.

2. Quels autres résultats positifs peut-on obtenir dans le sens indiqué par Hilbert, c'est-à-dire l'existence de "bonnes" fonctions de n variables non représentables au moyen d'un nombre fini de superpositions de "bonnes" fonctions de $n-1$ variables ? Une série de résultats dans cette direction est due au mathématicien soviétique A. G. Vitushkin ; nous en citerons quelques uns.

3. Dans la formule (3), qu'on appelle le théorème de superposition de Kolmogorov, peut-on imposer aux φ_{pq}, ou aux g_q, d'être de "bonnes" fonctions, dans un sens à préciser ? Evidemment, les réponses aux questions 2. et 3. constituent des sortes de réciproques les unes des autres.

4. Y-a-t-il une façon imagée de se représenter le théorème de superposition de Kolmogorov ? Quel rôle y joue l'intervalle $[1\ ,\ 2n+1]$ où varie l'entier q ?

5. Y-a-t-il des applications de ce théorème à d'autres domaines de l'analyse ?

6. Si l'on revient à la motivation première de Hilbert, il s'agit d'équations algébriques. Que devient le problème si on se restreint à des fonctions algébriques des variables ?

Mon intention n'est pas de faire exactement le point sur toutes ces questions. Le lecteur intéressé pourra se reporter à [1], [2], [3]. [1] est le compte-rendu d'un colloque organisé en 1974 sur les problèmes de Hilbert par l'American Mathematical Society ; il contient notamment l'énoncé du 13ème problème (pp. 20-21), une étude de G. G. Lorentz (pp. 419-430), et un commentaire de V. Arnold et G. Shimura (pp. 45-46) relatif à la question n. 6 (superposition de fonctions algébriques). [2] contient une foule d'informations, plusieurs démonstrations complètes, et des idées importantes ; il peut intéresser le lecteur de savoir que son auteur, Vitushkin, qui sait mener à merveille des calculs très compliqués, et semble connaître parfaitement toute la littérature sur le sujet, est un aveugle. [3] contient des réponses qui me sont personnelles, en particulier à la question n. 5.

Je vais plutôt tenter de justifier le titre que j'ai donné à cet article : un carrefour de l'algèbre, de l'analyse et de la géométrie. Je vais d'abord exposer de l'algèbre très ancienne : la théorie des équations algébriques suivant Tschirnhaus (1651-1708), qui explique pourquoi les équations de degré 6 peuvent être résolues au moyen de fonctions de deux variables, et pourquoi l'équation générale du 7ème degré se ramène à la forme (1) au moyen d'additions et de fonctions d'une variable[4]. Puis j'indiquerai quelques résultats dans la voie de Hilbert, et le rôle qu'y joue la notion d'objet générique - ou, de façon équivalente, de propriété quasi sûre - ; nous verrons aussi apparaître l'ε-entropie de Kolmogorov (1955)[5], qui explique et permet maintenant d'exposer sans trop de mal les résultats de Vitushkin (1954)[6]. La partie "géométrique" sera l'interprétation du théorème de superposition de Kolmogorov ; là encore, on voit apparaître des objets génériques, et un mélange plaisant d'analyse et de topologie. Enfin, je tenterai de conclure avec quelques réflexions sur ce qui paraît important ou intéressant en mathématiques, sur les facteurs de permanence et les facteurs d'évolution.

I. LES EQUATIONS ALGEBRIQUES ET LA TRANSFORMATION DE TSCHIRNHAUS.

Les fondateurs de la théorie des équations algébriques sont les italiens Niccolo Fontana, dit Tartaglia (1499-1557), qui donna la première formule pour résoudre les équations du 3ème degré, Gerolamo Cardano (1501-1576) qui la publia dans son "Ars Magna" (1545), Ludovico Ferrari (1522-1565) qui, dans l'"Ars Magna", donna la résolution d'équations de degré 4, et Raffaelo Bombelli (1522-1573) qui, dans son "Algebra" (1572), introduisit les racines imaginaires pour donner la solution générale des équations de degré 3 et 4.[7]

L'équation générale du 3ème degré se ramène, par translation de l'inconnue, à $X^3 + pX + q = 0$, qui admet pour solution

$$X = \left(-\frac{q}{2} + \left(\frac{q^2}{4} + \frac{p^3}{27}\right)^{1/2}\right)^{1/3} + \left(-\frac{q}{2} - \left(\frac{q^2}{4} + \frac{p^3}{27}\right)^{1/2}\right)^{1/3}$$

(formule de G. Cardano). On le voit en posant $X = u + v$ avec $u^3 + v^3 = -q$.

L'équation générale du 4ème degré peut se résoudre en superposant additions, multiplications, racines carrées, racines cubiques et racines quatrièmes. Il est inutile de tenter d'écrire une formule, qui serait très compliquée. Nous allons voir que la méthode de Tschirnhaus rend cela très facile.

Pour les équations de degré supérieur, on sait qu'elles ne sont pas solubles, en général, au moyen d'additions, multiplications, et radicaux. C'est le résultat principal de la théorie de Galois (19ème siècle !).

Avant Galois, et en particulier au cours du 17ème siècle, les tentatives n'avaient pas manqué pour tenter de résoudre "par radicaux" les équations de degré ≥ 5. L'idée de Tschirnhaus (1683) est la suivante.

Soit P_n un polynôme de degré n $(n \geq 3)$. On cherche à résoudre $P_n(X) = 0$. Posons $Y = Q_{n-1}(X)$, Q_{n-1} étant un polynôme de degré $n-1$ à déterminer. Eliminons X entre les équations $P_n(X) = 0$ et $Q_{n-1}(X) - Y = 0$. Le résultant de ces deux équations est un polynôme en Y, de degré n, qu'on peut essayer de résoudre ou au moins de simplifier par un bon choix du polynôme Q_{n-1}. Il restera alors à résoudre l'équation de degré $n-1$: $Q_{n-1}(X) - Y = 0$.

Essayons d'appliquer cette idée pour $n = 3, 4, 5, \ldots$; on écrira P et Q pour P_n et Q_{n-1}.

Pour $n = 3$, $P(X) = X^3 + aX^2 + bX + c$, $Q(X) = X^2 + \alpha X + \beta$. Le résultant de $P(X) = 0$ et $Q(X) - Y = 0$ est le déterminant

$$R(Y) = \begin{vmatrix} 1 & a & b & c & \\ & 1 & a & b & c \\ 1 & \alpha & \beta - Y & & \\ & 1 & \alpha & \beta - Y & \\ & & 1 & \alpha & \beta - Y \end{vmatrix}$$

que nous écrivons

$$R(Y) = -Y^3 + P_1(\alpha, \beta)Y^2 + P_2(\alpha, \beta)Y + P_3(\alpha, \beta).$$

Comme $P_1(\alpha, \beta)$ et $P_2(\alpha, \beta)$ sont respectivement des polynômes de degré 1 et 2 en α et β, il est tout à fait facile de calculer α et β pour avoir $P_1(\alpha, \beta) = P_2(\alpha, \beta) = 0$, et alors la résolution de $R(Y) = 0$ et $Q_n(X) - Y = 0$ est immédiate. Comme exercice, le lecteur peut essayer d'obtenir à partir de là la formule de G. Cardano.

Pour $n = 4$, $P(X) = X^4 + aX^3 + bX^2 + cX + d$, $Q(X) = X^3 + \alpha X^2 + \beta X + \gamma$, et

$$R(Y) = Y^4 + P_1(\alpha,\beta,\gamma)\,Y^3 + P_2(\alpha,\beta,\gamma)\,Y^2 + P_3(\alpha,\beta,\gamma)\,Y + P_4(\alpha,\beta,\gamma)\,,$$

où $P_j(\alpha,\beta,\gamma)$ désigne un polynôme de degré j en α,β,γ. Quitte à résoudre une équation du 3ème degré, on peut choisir α,β,γ de façon que $P_1(\alpha,\beta,\gamma) = P_3(\alpha,\beta,\gamma) = 0$ (on a une grande liberté de choix). Alors $R(Y) = 0$ est une équation bicarrée, qu'on sait résoudre, et il reste ensuite à résoudre l'équation de 3ème degré $Q(X) - Y = 0$.

Pour $n \geq 5$, $P(X) = X^n + a_1X^{n-1} + a_2X^{n-2} + \ldots + a_n\,,$

$$Q(X) = X^{n-1} + \alpha_1X^{n-2} + \ldots + \alpha_{n-1}\,,$$

$$R(Y) = (-1)^n\,Y^n + \sum_{m=1}^{n} P_m(\alpha_1,\ldots,\alpha_{n-1})\,Y^{n-m}\,.$$

Quand on fixe $\alpha_4,\ldots,\alpha_{n-1}$, les équations $P_1(\alpha_1,\ldots,\alpha_{n-1}) = 0$, $P_2(\alpha_1,\ldots,\alpha_{n-1}) = 0$ et $P_3(\alpha_1,\ldots,\alpha_{n-1}) = 0$ représentent respectivement un plan, une surface du second degré et une surface du 3ème degré dans $\mathbb{R}^3$ (où les coordonnées sont α_1, α_2, α_3). Par un choix convenable de $\alpha_4,\ldots,\alpha_{n-1}$, se ramenant à la résolution d'équations de degré ≤ 3, le plan $P_1 = 0$ et la quadrique $P_2 = 0$ sont tangents. Leur intersection est constituée de deux droites. L'intersection de ces droites avec la surface $P_3 = 0$ donne les points communs à $P_1 = 0$, $P_2 = 0$, $P_3 = 0$. On vérifie que $\alpha_1,\ldots,\alpha_{n-1}$ s'expriment à partir de $a_1, a_2, \ldots, a_n$ au moyen des opérations de l'arithmétique (additions, multiplications, divisions) et de racines carrées et cubiques. Ainsi la résolution de l'équation $P(X) = 0$ se ramène à la résolution d'une équation

$$Y^n + b_4\,Y^{n-4} + b_5\,Y^{n-5} + \ldots + b_n = 0 \tag{4}$$

suivie de la résolution de l'équation $Q(X) - Y = 0$. Quitte à prendre pour inconnue $Z = Y\,b_n^{-1/n}$, on peut supposer $b_n = 1$ dans (4). Ainsi la solution générale de l'équation du 5ème degré s'exprime à l'aide des opérations de l'arithmétique, des radicaux, et de la fonction $X = \varphi(x)$ solution de l'équation

$$X^5 + xX + 1 = 0\,. \tag{5}$$

La solution générale de l'équation du 6ème degré s'exprime à l'aide des opérations de l'arithmétique, des radicaux, de la fonction $\varphi(x)$, et de la fonction $\psi(x,y)$ solution de l'équation

$$X^6 + xX^2 + yX + 1 = 0. \tag{6}$$

Enfin - et c'est là le point de départ de Hilbert - la solution générale de l'équation du 7ème degré s'exprime à l'aide des opérations de l'arithmétique, des radicaux, de la fonction $\varphi(x)$, de la fonction $\psi(x,y)$, et de la fonction $\chi(x,y,z)$ solution de l'équation (1).

Peut-on exprimer la fonction algébrique $\chi(x,y,z)$ au moyen des superpositions de fonctions algébriques de deux variables ? La question est toujours ouverte.

Voici une autre question ouverte, à la fois plus générale et plus précise : peut-on exprimer toute fonction algébrique, d'un nombre quelconque de variables, au moyen de superpositions de fonctions algébriques d'une variable, d'additions, de multiplications et de divisions ?

Rappelons que par définition une fonction algébrique de n variables $x_1, x_2, \ldots, x_n$ est la solution $f = f(x_1, x_2, \ldots, x_n)$ d'une équation algébrique $p_o f^m + p_1 f^{m-1} + \ldots + p_m = 0$ dont les coefficients p_j sont des polynômes en $x_1, x_2, \ldots, x_n$. C'est, en général - comme dans le cas des radicaux - une fonction multiforme. Si $p_o = 1$, on dit que f est une fonction algébrique entière. Voici un résultat curieux de A. G. Hovanski (1971) (voir [2] pour la référence) : la solution de l'équation

$$X^5 + xX^2 + yX + 1 = 0$$

n'est pas représentable au moyen de fonctions algébriques entières d'une variable, d'additions et de multiplications. Cela ne contredit pas le résultat que nous avons établi à l'aide de la transformation de Tschirnhaus, à savoir que la solution générale de l'équation de 5ème degré s'exprime à l'aide des opérations de l'arithmétique, des radicaux, et de la fonction algébrique entière $\varphi(x)$ solution de (5). Cela montre simplement que, parmi les opérations de l'arithmétique, il ne faut pas oublier la division !

II. DANS LA VOIE DE HILBERT ET DE VITUSHKIN.

On peut composer des fonctions, ou substituer des fonctions à des variables, ou superposer des fonctions. Nous avons déjà utilisé ces expressions et, dans cet article, elles sont synomymes. En relation avec le 13ème problème de Hilbert, le terme le plus courant est celui de superposition, sous l'influence de l'école russe.

Précisons ce qu'on entend par superposition d'ordre s de fonctions de p variables appartenant à une classe $\mathcal{C}$, opérant sur $x_1, x_2, \dots, x_n$. Les superpositions d'ordre 0 sont les fonctions $x_1, x_2, \dots, x_n$. Si $U_1, U_2, \dots, U_p$ sont des superpositions d'ordre $s-1$ et si $\varphi(t_1, t_2, \dots, t_p)$ est une fonction de la classe $\mathcal{C}$, $\varphi(U_1, U_2, \dots, U_p)$ est une superposition d'ordre s. La définition se transcrit immédiatement pour des germes de fonctions ou pour des séries formelles.

Considérons un exemple simple. Prenons pour $\mathcal{C}$ la classe des polynômes de deux variables, à coefficients réels (ou complexes si on préfère), opérant sur x_1, x_2, x_3. Regardons les superpositions d'ordre s, modulo les monômes de degré strictement supérieur à m. Elles sont identifiables à des polynômes de degré $\le m$ en x_1, x_2, x_3, et elles dépendent d'un nombre de paramètres que nous désignerons par $\gamma(m,s)$. Ainsi $\gamma(0,s) = 1$ et $\gamma(1,s) = 4$ quel que soit $s \ge 1$. Le calcul de $\gamma(m,1)$ est un exercice très simple : pour $k \ge 1$, il existe $3k$ monômes de degré k en x_1, x_2, x_3 où figurent seulement deux des variables, donc

$$\gamma(m,1) = 1 + 3 + 3 \times 2 + \dots + 3m = 1 + 3\,\frac{m(m+1)}{2}.$$

Pour $\gamma(m,s)$, nous allons chercher une majoration. Soit $\varphi \in \mathcal{C}$, et soient U_1 et U_2 deux superpositions arbitraires d'ordre $s-1$. Modulo les monômes de degré $> m$, U_1 dépend de $\gamma(m, s-1)$ paramètres, et de même pour U_2. Modulo les monômes de degré $> m$, $\varphi(t_1,t_2)$ dépend de $1 + 2 + 3 + \dots (m+1)$ paramètres. Donc, modulo les monômes de degré $> m$, $\varphi(U_1,U_2)$ dépend d'au plus $\frac{(m+1)(m+2)}{2} + 2\gamma(m,s-1)$ paramètres. Ainsi

$$\gamma(m,s) \le \frac{(m+1)(m+2)}{2} + 2\gamma(m, s-1)$$

d'où

$$\gamma(m,s) + \frac{(m+1)(m+2)}{2} \le 2(\gamma(m,s-1) + \frac{(m+1)(m+2)}{2}) \le \ldots\ldots$$

$$\le 2^{s-1}\,(\gamma(m,1) + \frac{(m+1)(m+2)}{2})$$

et enfin

$$\gamma(m,s) \le 2^s(m+2)^2.$$

Or les polynômes de degré $\le m$ en x_1, x_2, x_3, dépendent de d_m paramètres, avec

$$d_m = \frac{(m+1)^3}{6} + \frac{(m+1)^2}{2} + \frac{m+1}{3}.$$

Pour chaque s fixé, on voit que $d_m > \gamma(m,s)$ lorsque m est assez grand. Exprimons cela d'autre manière. Les polynômes de degré $\le m$ en x_1, x_2, x_3 forment une variété $\mathcal{P}_m$ de dimension d_m. Les sections de degré m (c'est-à-dire la somme des monômes de degré $\le m$) dans les superpositions d'ordre s de polynômes à deux variables forment une sous-variété $Q_{m,s}$ de $\mathcal{P}_m$, de dimension $\gamma(m,s)$, et, pour m assez grand (disons $m \ge m_s$), $Q_{m,s} \ne \mathcal{P}_m$.

Que peut-on tirer de cette remarque ? Rien d'intéressant pour les polynômes, puisqu'évidemment tout polynôme en x_1, x_2, x_3 s'écrit comme superposition de polynômes de deux variables.

Par contre, prenons maintenant pour $\mathcal{C}$ la classe des séries formelles de deux variables

$$\varphi(t_1,t_2) = \sum_{(j,k)\in\mathbb{N}^2} a_{jk}\, t_1^j\, t_2^k .$$

A partir de x_1, x_2, x_3, en superposant des séries formelles de deux variables, on obtient des séries formelles de trois variables

(7) $$\sigma(x_1,x_2,x_3) = \sum_{(p,q,r)\in\mathbb{N}^3} c_{pqr}\, x_1^p\, x_2^q\, x_3^r .$$

Soit $\mathcal{S}$ l'ensemble de toutes les séries formelles en x_1, x_2, x_3, soit Q_s l'ensemble de toutes celles qui s'obtiennent comme superpositions d'ordre s de séries formelles à deux variables, et soit $Q = \bigcup_s Q_s$. Evidemment $Q \subset \mathcal{S}$. Je dis que $Q \ne \mathcal{S}$, c'est-à-dire qu'il existe des séries formelles de trois variables non représentables comme superpositions de séries formelles de deux variables.

En effet, désignons par $\sigma_m(x_1,x_2,x_3)$ la section de degré m de la série (7), c'est-à-dire le polynôme

$$\sigma_m(x_1,x_2,x_3) = \sum_{0\leq p+q+r\leq m} c_{pqr}\, x_1^p\, x_2^q\, x_3^r \,.$$

On a $\sigma_m \in \mathcal{P}_m$. Les sections de degré m des superpositions d'ordre s de séries formelles à deux variables forment la sous-variété de $\mathcal{P}_m$ que nous avons appelée $Q_{m,s}$. Pour $m = m_s$, $Q_{m,s}$ est une sous-variété propre de $\mathcal{P}_m$, donc $\mathcal{P}_m \setminus Q_{m,s}$ est un ouvert dense de $\mathcal{P}_m$. Pour la topologie naturelle $\mathcal{T}$ (celle de la convergence simple des coefficients c_{pqr}), l'ensemble des $\sigma \in \mathcal{T}$ tels que $\sigma_{m_s} \in \mathcal{P}_{m_s} \setminus Q_{m_s,s}$ est un ouvert dense de $\mathcal{T}$, soit $\mathcal{G}_s$. Evidemment $\mathcal{G}_s \cap Q_s = \emptyset$. Par sélection des coefficients, il est facile de construire une série formelle $\sigma \in \mathcal{T}$ telle que $\sigma \in \mathcal{G}_s$ pour tout s. Alors $\sigma \notin Q = \bigcup_s Q_s$, ce qui prouve mon assertion.

En fait, nous avons prouvé ceci : l'ensemble des $\sigma \in Q$ contient une intersection dénombrable d'ouverts dense de $\mathcal{T}$. Dans un espace métrisable complet (c'est le cas de $\mathcal{T}$) on appelle ensemble de première catégorie au sens de Baire tout ensemble contenu dans une réunion dénombrable d'ensembles fermés dont le complémentaire est dense, et ensemble de deuxième catégorie tout ensemble contenant une intersection dénombrables d'ouverts denses. Un ensemble de deuxième catégorie est le complémentaire d'un ensemble de première catégorie. On établit très facilement - c'est le "théorème de Baire" - qu'un ensemble de deuxième catégorie est dense dans l'espace considéré. Les ensembles de première catégorie doivent être regardés comme maigres (ils jouent un rôle analogue aux ensembles de mesure nulle dans la théorie de la mesure) et les ensembles de seconde catégorie comme gros (leurs analogues en théorie de la mesure sont les ensembles de mesure pleine). On dit qu'une propriété est générique si elle a lieu sur un ensemble de seconde catégorie. On dit aussi "quasi tout point" pour signifier "tout point dans un certain ensemble de seconde catégorie". Ainsi, pour les séries formelles de trois variables, la propriété de ne pas être représentable comme superposition de séries formelles de deux variables est une propriété générique ; elle a lieu pour quasi toute série formelle de trois variables.

Au prix d'un petit effort supplémentaire, on obtient assez facilement l'énoncé suivant.

THEOREME. Quasi toute fonction entière $F(x_1,x_2,x_3)$ de trois variables a la propriété suivante : quel que soit $\xi = (\xi_1,\xi_2,\xi_3) \in \mathbb{R}^3$, le germe de F en ξ ne peut pas s'obtenir par superposition de séries formelles de deux variables.

C'est vraisemblablement ce résultat - exprimé de façon constructive et non comme application du théorème de Baire, qui n'était pas encore énoncé[8]- que Hilbert avait à l'esprit en écrivant qu'il savait construire une fonction analytique de trois variables non réalisable par une chaîne finie de substitutions au moyen de fonctions de deux variables. Dans l'énoncé du théorème, "quasi toute fonction entière" s'entend naturellement au sens de la topologie usuelle sur les fonctions entières, mais on peut, avec beaucoup de liberté, changer d'espace et de topologie. On peut aussi écrire le théorème pour des fonctions de n variables $F(x_1,\ldots,x_n)$. Il en résulte l'existence de fonctions entières $F(x_1,x_2,\ldots,x_n)$, non représentables comme superpositions de fonctions de $n-1$ variables de classe C^∞. Toutes ces variantes sont faciles et n'exigent pas d'idée nouvelle.

La signification du théorème de Hilbert est que les fonctions analytiques de n variables forment un espace beaucoup plus riche, plus complexe que les fonctions analytiques de $n-1$ variables. Pour les fonctions analytiques, le nombre de variables est donc, dans un certain sens, une mesure de la complexité de l'espace. Il en est de même, on vient de le remarquer, pour les fonctions de classe C^∞.

Considérons maintenant les fonctions réelles de classe C^α de n variables réelles $(n = 2, 3, \ldots)$; $\alpha > 0$; si α est entier, on dit que $f \in C^\alpha$ si f admet des dérivées partielles continues Df jusqu'à l'ordre α ; si $\alpha = \nu + \beta$, ν entier, $0 < \beta < 1$, on dit que $f \in C^\alpha$ si $f \in C^\nu$ et que chaque dérivée partielle Df d'ordre $\leq \nu$, sur chaque boule de R^n, satisfait une condition du type

$$|Df(x) - Df(x')| \leq C\|x - x'\|^\beta .$$

Ces fonctions constituent un espace $C^{\alpha}(R^n)$, que nous notons plus rapidement C_n^{α}. On va voir qu'ici c'est le nombre $\frac{n}{\alpha}$ qui permet de mesurer la complexité de l'espace. Voici en effet un théorème de A. G. Vitushkin (1954).

THEOREME. Supposons $\frac{n}{\alpha} > \frac{n'}{\alpha'}$ (inégalité stricte), et $\alpha' \geq 1$. Alors quasi toute fonction dans C_n^{α} est non représentable par superpositions de fonctions de la classe $C_{n'}^{\alpha'}$ (opérant sur $x_1, x_2, \ldots, x_n$).

La meilleure démonstration de ce théorème est fournie par la notion d'ε-entropie, due à Kolmogorov (1955). Voici de quoi il s'agit. Soit $\mathcal{E}$ un espace métrique(1), et E un sous-ensemble de $\mathcal{E}$ tel que, pour tout $\varepsilon > 0$, E puisse être recouvert par un nombre fini de boules de rayon ε. Soit $N(\varepsilon) = N(\varepsilon, E, \mathcal{E})$ le nombre minimum de ces boules. On appelle ε-entropie de E dans $\mathcal{E}$ la fonction

$$H(\varepsilon) = H(\varepsilon, E, \mathcal{E}) = \log N(\varepsilon) \qquad (\varepsilon > 0).$$

Si $\mathcal{E} = \mathbb{R}^n$ et si E est une sphère de dimension $p < n$ plongée dans $\mathbb{R}^n$, on voit très facilement que $H(\varepsilon) \sim p \log \frac{1}{\varepsilon}$. Ainsi la notion d'$\varepsilon$-entropie peut être utilisée pour définir une sorte de dimension pour certains objets géométriques. Mais elle est surtout utile en analyse fonctionnelle, quand $\mathcal{E}$ est un espace fonctionnel.

Prenons pour $\mathcal{E}$ l'espace $C_0(\mathbb{R}^n)$ constitué par les fonctions continues sur $\mathbb{R}^n$ qui tendent vers 0 à l'infini. C'est un espace vectoriel normé, la norme de f étant $\|f\|_0 = \sup_x |f(x)|$ $(x \in \mathbb{R}^n)$. Soit $\mathfrak{F}$ l'ensemble des fonctions g de la classe C_n^{α} qui s'annulent en dehors d'une boule de $\mathbb{R}^n$ (dépendant de g) ; on définit la norme par

$$\|g\|_{\alpha} = \Sigma \|D^p g\|_0 + \sup_{x,y} \left| \frac{D^p g(x) - D^p g(y)}{\|x-y\|^{\alpha}} \right|$$

la somme étant prise sur tous les opérateurs de dérivation D^p d'ordre inférieur à α (on suppose que α n'est pas entier). Prenons pour E une boule de $\mathfrak{F}$. Le lemme essentiel, pour le théorème de Vitushkin, s'exprime par la formule simple

$$H(\varepsilon, E, \mathcal{E}) \approx \left(\frac{1}{\varepsilon}\right)^{\frac{n}{\alpha}},$$

le signe $\approx$ signifiant que le rapport des deux membres est compris entre deux nombres strictement positifs indépendants de ε. Ainsi $\frac{n}{\alpha}$, comme je l'ai déjà dit, est une mesure de la complexité de l'espace C_n^α.

A partir de là, l'idée de la démonstration du théorème est assez simple, même si les détails sont compliqués. On part de $C_{n'}^{\alpha'}$, $\mathfrak{F}'$ et E' ($\mathfrak{F}'$ et E' étant définis à partir de $C_{n'}^{\alpha'}$ comme $\mathfrak{F}$ et E à partir de C_n^α). On vérifie qu'en superposant des fonctions appartenant à E', on ne change pas l'ordre de grandeur de l'ε-entropie. Comme on a supposé $\frac{n}{\alpha} > \frac{n'}{\alpha'}$, les superpositions de fonctions appartenant à E' (et aussi les superpositions de fonctions appartenant à $C_{n'}^{\alpha'}$) forment un ensemble trop maigre pour recouvrir C_n^α.

Le théorème de Vitushkin de 1954 est très intéressant en lui-même, comme une réponse partielle au problème de Hilbert. Mais n'est-il pas vrai que sa démonstration à l'aide de l'ε-entropie est plus intéressante encore ? Avec l'ε-entropie, on voit se dégager nettement une notion générale permettant de traiter de la complexité des espaces fonctionnels.

Voici un autre théorème de Vitushkin (1964),[9] difficile, et dont l'importance va mieux apparaître après nos commentaires sur le théorème de Kolmogorov (1957).

THEOREME. <u>Il existe une fonction analytique</u> $F(x_1,x_2)$ <u>qui n'est pas représentable sous la forme</u>

$$\sum_{i=1}^{N} p_i(x_1,x_2)\,\varphi_i(q_i(x_1,x_2))$$

<u>lorsque les</u> p_i et q_i <u>sont des fonctions fixées de deux variables de classe</u> C^1 <u>et les</u> φ_i <u>des fonctions continues arbitraires</u>.

---

(1) Dans un espace métrique on sait définir la distance entre deux points, et, en conséquence, on sait définir une boule centrée en un point donné et ayant un rayon donné. Un ouvert, dans un espace métrique, est une réunion de boules centrées en chacun de ses points ; un fermé est le complémentaire d'un ouvert. L'intérieur d'un ensemble A est le plus grand ouvert contenu dans A ; la fermeture d'un ensemble B est le plus petit fermé contenant B. Un ensemble est dense si sa fermeture est tout l'espace. Un ensemble maigre, au sens de Bourbaki, est un ensemble contenu dans une réunion dénombrable de fermés d'intérieur vide ; c'est exactement un ensemble de 1ère catégorie au sens de Baire. Dans un espace vectoriel normé la distance entre deux points x et y est $\|x-y\|$.

III. COMMENT "VOIR" LE THEOREME DE KOLMOGOROV ?

Rappelons le théorème, qui a été donné dans l'introduction (formule (3)). Soit I l'intervalle $[0,1]$ de $\mathbb{R}$. On appelle I^n le cube unité de $\mathbb{R}^n$, et $C(I^n)$ l'espace des fonctions à valeurs réelles continues sur I^n.

THEOREME. Toute $f \in C(I^n)$ $(n \geq 2)$ peut s'écrire sous la forme

(8) $$f(x_1, x_2, \ldots, x_n) = \sum_{q=1}^{2n+1} g_q \left(\sum_{p=1}^{n} \varphi_{pq}(x_p) \right),$$

où les g_q et les φ_{pq} sont des fonctions continues d'une variable réelle, les φ_{pq} sont strictement croissantes sur I, et les φ_{pq} ne dépendent pas de f.

On peut "voir" ce théorème en se plaçant dans l'espace $\mathbb{R}^{2n+1}$. Pour chaque p $(p = 1, 2, \ldots, n)$, considérons l'arc de courbe Γ_p défini par les équations

(9) $$X_q = \varphi_{pq}(t) \qquad (t \in I,\ q = 1, 2, \ldots, 2n+1).$$

Comme toutes les coordonnées sont des fonctions croissantes de t, nous dirons que c'est un "arc croissant".

Considérons maintenant la somme algébrique

(10) $$E = \Gamma_1 + \Gamma_2 + \ldots + \Gamma_p$$

c'est-à-dire l'ensemble des points de $\mathbb{R}^{2n+1}$ qui s'obtiennent en ajoutant un point de Γ_1, un point de Γ_2, ... un point de Γ_p. C'est l'image du cube I^n par l'application $(x_1, x_2, \ldots, x_n) \rightsquigarrow (X_1, X_2, \ldots, X_{2n+1})$ défini par

$$X_q = \sum_{p=1}^{n} \varphi_{pq}(x_p) \qquad (q = 1, 2, \ldots, 2n+1).$$

Que dit le théorème ?

A) E est une image bijective de I^n, c'est-à-dire qu'à deux points différents de I^n correspondent deux points différents de E. Au point de vue topologique, E n'est pas autre chose que le cube I^n, tordu et placé dans $\mathbb{R}^{2n+1}$ de façon que les arêtes issues de 0 se transforment en les arcs de courbe Γ_p.

B) Toute fonction continue sur E peut s'écrire

(11) $$g_1(X_1) = g_2(X_2) + \ldots + g_{2n+1}(X_{2n+1}) \qquad (g_q \in C(\mathbb{R})).$$

Ainsi E est un "ensemble d'interpolation" dans le sens qu'a ce terme dans plusieurs branches de l'analyse : E est une partie de $\mathbb{R}^{2n+1}$, et toute fonction $F \in C(E)$ (continue sur E) est prolongeable en une fonction $G \in C(\mathbb{R}^{2n+1})$ d'une forme bien particulière, à savoir (11).

Depuis Kolmogorov, des améliorations de ce théorème ont été cherchées dans deux directions : quelles conditions supplémentaires peut-on imposer aux fonctions φ_{pq} ? Quelles conditions supplémentaires peut-on imposer aux fonctions g_q ? Je vais en dire quelques mots.

Quoique l'énoncé du théorème soit paradoxal, il traduit une propriété générique des φ_{pq}.

THEOREME. <u>Pour quasi tout choix des</u> φ_{pq}, <u>fonctions continues croissantes définies sur</u> I, <u>l'ensemble</u> E <u>a les propriétés</u> A) <u>et</u> B).

En fait, il est un peu plus facile de démontrer ce théorème de Kolmogorov générique que de construire explicitement les φ_{pq}. Les lecteurs intéressés à la démonstration peuvent la trouver dans [2].

Certainement on ne peut pas choisir les φ_{pq} de classe C^1. C'est une conséquence du théorème de Vitushkin (1964) signalé à la fin de la partie II. Peut-on les choisir dans une classe Λ_α (classe de Lipschitz-Hölder d'ordre α) $(0 < \alpha \leq 1)$ définie comme l'ensemble des φ tels que

$$|\varphi(t') - \varphi(t)| \leq C_\varphi |t - t'|^\alpha \quad ?$$

La réponse est positive. On le voit facilement, en étudiant de près la construction de Kolmogorov, pour $0 < \alpha < 1$. Pour $\alpha = 1$, le problème paraissait beaucoup plus difficile, puisque la classe Λ_1 est assez voisine de la classe C^1, et la première solution donnée (1967) était compliquée[10]; la raison de cette complication me paraît être le fait que le théorème de Kolmogorov <u>générique</u> est vrai pour les $\varphi_{pq} \in \Lambda_\alpha$, faux pour les $\varphi_{pq} \in \Lambda_1$.

Mais le fait que le théorème de Kolmogorov ordinaire soit vrai avec la restriction $\varphi_{pq} \in \Lambda_1$ est néanmoins presque évident. Il suffit de songer à l'interprétation géométrique. Une fois définis les arcs $\Gamma_1, \dots, \Gamma_n$, les propriétés A) et B) ne dépendent plus du paramétrage. Comme les arcs Γ_p sont "croissants", ils sont rectifiables, et on peut les paramétrer à l'aide de la longueur d'arc s, normalisée de façon que la longueur de chaque Γ_p soit 1. A la place de (9), les équations des Γ_p sont maintenant

$$X_q = \psi_{pq}(s) \qquad (s \in I, \quad q = 1, 2, \dots, 2n+1)$$

et les ψ_{pq} appartiennent à Λ_1. La formule (8) peut s'écrire avec les ψ_{pq} au lieu des φ_{pq}. Autrement dit, on peut imposer la condition $\varphi_{pq} \in \Lambda_1$.

On peut d'autre part diminuer le nombre de fonctions φ_{pq}, de la manière suivante (D. A. Sprecher 1964)[11]. Soit $\lambda_1, \lambda_2, \dots, \lambda_n$ des nombres strictement positifs tels que $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$, et tels que les 2^n sommes $\varepsilon_1\lambda_1 + \varepsilon_2\lambda_2 + \dots + \varepsilon_n\lambda_n$ ($\varepsilon_p = 0$ ou 1) soient toutes distinctes. Le théorème de Kolmogorov ordinaire et le théorème de Kolmogorov générique restent vrais en prenant $\varphi_{pq}(t) = \lambda_p\, \varphi_q(t)$, où les φ_q $(q = 1, 2, \dots, 2n+1)$ sont des applications continues croissantes de I sur I. Ainsi, en désignant par Γ l'arc

$$X_q = \varphi_q(t) \qquad (t \in I, \quad q = 1, 2, \dots, 2n+1)$$

qui joint les points $(0, 0, \dots, 0)$ et $(1, 1, \dots, 1)$ de $\mathbb{R}^{2n+1}$, on a $\Gamma_p = \lambda_p \Gamma$, et

$$E = \lambda_1\Gamma + \lambda_2\Gamma + \dots + \lambda_n\Gamma \,. \tag{12}$$

E est inscrit dans I^{2n+1}, et contient les sommets opposés $(0, 0, \dots, 0)$ et $(1, 1, \dots, 1)$. Les propriétés A) et B) de E sont des propriétés génériques de Γ.

Dans cette version, seules importent les valeurs des g_q sur I, donc on peut écrire dans (11) $g_q \in C(I)$ au lieu de $g_q \in C(\mathbb{R})$. Une intéressante remarque, due à G. G. Lorentz,[12] est qu'on peut les prendre tous égaux : $g_1 = g_2 = \ldots = g_{2n+1} = g$. Ainsi, à chaque $f \in C(I^n)$ on peut faire correspondre <u>une</u> fonction $g \in C(I)$.

Une seconde remarque, qui provient tout de suite de la démonstration, est qu'au lieu de prendre f et les g_p à valeurs réelles, on peut leur donner comme domaine de valeurs un espace fonctionnel. Il en résulte que, dans l'énoncé de Kolmogorov, on peut imposer aux g_p (ou à l'unique fonction g) de dépendre continûment de f.

Enfin, au lieu de choisir $g \in C(I)$, on peut imposer à g des conditions supplémentaires ; par exemple certaines propriétés de son développement de Fourier. Sans donner les détails, j'en dirai un mot tout à l'heure.

Le choix des g_q est intéressant pour la partie B) du théorème de Kolmogorov, c'est-à-dire les propriétés de l'ensemble E comme ensemble d'interpolation. Laissons cela de côté pour considérer seulement la condition A), qui veut dire que la somme $\Gamma_1 + \Gamma_2 + \ldots + \Gamma_n$ est directe, ou - de manière équivalente - que E n'a pas de point double.

Il est naturel de regarder la question suivante. On considère n arcs croissants $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ dans $\mathbb{R}^m$ $(m \geq n)$, et la somme algébrique

$$E = \Gamma_1 + \Gamma_2 + \ldots + \Gamma_n.$$

Est-il vrai que, génériquement, E n'ait pas de point double ? Cela dépend de la valeur de m.

Si $m \geq 2n+1$, génériquement, E n'a pas de point double. C'est la première partie du théorème de Kolmogorov générique.

Si $m \leq 2n$, génériquement, E a des points doubles. La démonstration est un bon exercice de topologie [3].

On voit ainsi que, dans le théorème de Kolmogorov générique, on ne peut pas remplacer $\mathbb{R}^{2n+1}$ par $\mathbb{R}^{2n}$. Pour le théorème de Kolmogorov ordinaire, à ma connaissance, le problème est toujours ouvert. Dans le cas $n = 2$, il a été résolu par R. Doss (1963).[13]

Je crois que, par son aspect géométrique, le théorème de Kolmogorov peut être séduisant pour beaucoup de mathématiciens. Je me permettrai d'indiquer ma motivation personnelle. En 1966-67, j'ai donné un cours à Orsay sur l'approximation des fonctions, en utilisant l'excellent livre de G.-G. Lorentz, Approximation of functions, qui venait de paraître. En analyse de Fourier, je connaissais la puissance de la méthode de Baire. J'étais d'autre part intéressé - et je le suis toujours - à un problème de Lusin qui s'énonce ainsi[14] : étant donné une fonction $f(t)$ continue sur le cercle $(t \in \mathbf{T} = \mathbf{R}/\mathbf{Z})$, peut-on trouver un homéomorphisme h de $\mathbf{T}$ (c'est-à-dire un changement de variable continu) tel que la fonction composée $f(h(t))$ ait sa série de Fourier absolument convergente (on écrit alors $f(h(t)) \in A(\mathbf{T})$) ? En étudiant la démonstration du théorème de Kolmogorov, j'ai vu qu'il était générique, et que, dans le cas $n = 1$, il donnait une réponse partielle au problème de Lusin : toute fonction $f \in C(\mathbf{T})$ peut s'écrire

$$f(t) = g(\varphi_1(t)) + g(\varphi_2(t)) + g(\varphi_3(t))$$

avec $g \in A(\mathbf{T})$, et cela pour quasi tout choix des homéomorphismes $\varphi_1, \varphi_2, \varphi_3$. Dans ce cas l'ensemble E est la courbe d'équation

$$\theta_1 = \varphi_1(t) \qquad \theta_2 = \varphi_2(t) \qquad \theta_3 = \varphi_3(t) \ ;$$

c'est un ensemble d'interpolation pour $A(\mathbf{T}^3)$, c'est-à-dire que toute fonction $F \in C(E)$ peut se prolonger sur $\mathbf{T}^3$ en une fonction de la classe $A(\mathbf{T}^3)$. Il y a d'autres interprétations de la formule (8) dans le même esprit.

IV. POURQUOI S'INTERESSER AU 13ème PROBLEME DE HILBERT ?

A chaque époque, il y a de grands problèmes qui sont comme un défi posé aux mathématiciens. C'est comme un grand concours, ouvert quelquefois pendant des siècles, et où le vainqueur triomphe non seulement de ses contemporains, mais des grands hommes du passé. Celui qui résout un vieux problème, auquel des mathématiciens importants se sont attaqués, sait qu'il a fait quelque chose de nouveau, qu'il a marqué un jalon dans le progrès de la science. Aux côtés du problème de Fermat, de l'hypothèse de Riemann, de la conjecture de Poincaré, etc..., chacun des problèmes de Hilbert a joué - et parfois continue à jouer - ce rôle de défi, et les solutions, quand elles existent, constituent de bons jalons pour mesurer le chemin parcouru.

Noble discours. Mais regardons-y de plus près. Le problème de la duplication du cube, celui de la trisection de l'angle, celui de la quadrature du cercle, ont été des problèmes majeurs de la mathématique pendant plus de deux millénaires - Platon déjà proposait qu'on mette le premier au concours -. Chacun sait qu'ils ont une solution négative. Mais qui la connait ? Et parmi ceux qui la connaissent, qui connait le nom de Pierre-Laurent Wantzel, qui a résolu les deux premiers en 1837 (dans ses "recherches sur les moyens de reconnaître si un problème de géométrie peut se résoudre par la régle et le compas"), et celui de C. L. Ferdinand von Lindemann, qui a démontré la transcendance de π en 1882 [15] ? N'y a-t-il pas d'autres exemples, nombreux de problèmes oubliés en même temps que leurs solutions ? N'est-ce pas même le sort général d'un problème de ne plus servir à rien et d'être délaissé une fois résolu ?

Alors, reprenons la question : pourquoi s'intéresser au 13ème problème de Hilbert ?

- Sans doute, me direz-vous, parce que Hilbert avait de bonnes raisons de le poser. Si ces raisons étaient bonnes pour lui, elles restent sûrement bonnes pour nous.

- Alors regardons les raisons de Hilbert. Il cite au début le traité de nomographie de Maurice d'Ocagne (1899). Savez-vous ce que c'est que la nomographie ?

- Non.

- C'est un moyen de résoudre une équation en traçant une famille de courbes dépendant d'un paramètre. Hilbert nous explique en 30 lignes pourquoi les équations jusqu'au degré 6 peuvent être résolues au moyen de tables nomographiques convenables, et ensuite exprime son opinion que l'équation du 7ème degré n'est pas résoluble au moyen des constructions nomographiques. Dites-moi, qui se soucie maintenant de la solution des équations par tables ou constructions nomographiques ?

- Personne évidemment.

- Donc la motivation de Hilbert ne nous intéresse plus.

Arrêtons là le badinage. Même si les motivations de Hilbert ne sont plus les nôtres, il a posé un vrai problème. La preuve en est que, même résolu, ce problème nous intéresse. Il est d'autant plus intéressant que, précisément, des motivations très diverses peuvent y mener et en provenir. L'aspect algébrique est loin d'être épuisé - et c'est même sans doute le plus prometteur actuellement -. L'aspect analyse-géométrie-topologie-théorie de la dimension renoue avec une grande tradition. En 1877, G. Cantor découvrait l'existence d'une bijection entre l'intervalle $[0,1]$ et le carré $[0,1]\times[0,1]$; il en était bouleversé, et écrivait à R. Dedekind que cela ruinait l'idée qu'il se faisait de la dimension des variétés. R. Dedekind lui répondait, presque immédiatement, que le concept de dimension n'était nullement ruiné par sa découverte, et il indiquait comme problème à résoudre l'impossibilité d'une bijection continue entre deux variétés de dimensions différentes.[16] Le programme de Dedekind constitue l'acte de naissance de la théorie de la dimension. De façon analogue, les découvertes d'Arnold et de Kolmogorov heurtent l'intuition, mais rendent d'autant plus intéressantes les réponses positives apportées à la mesure de la complexité des espaces de fonctions. Et ces mesures de complexité interviennent actuellement dans tous les domaines mathématiques, y compris le codage des signaux (A. G. Vitushkin a fait là-dessus une des conférences majeures du Congrès International des Mathématiciens en 1974).[17]

Le lecteur est maintenant en droit de m'interroger.

- Ainsi, selon vous, un sujet mathématique est intéressant parce que les mathématiciens s'y intéressent. N'avez-vous vraiment pas d'autre explication à proposer ?

- Si, ami lecteur. Même si cela choque d'éminents collègues, il y a des explications à trouver en dehors des mathématiques. Voici, par exemple, une idée simple : à chaque époque, plusieurs concepts fondamentaux des mathématiques sont en relation directe avec les moyens de calcul dont on dispose à cette époque. La théorie des fonctions du 17ème au 19ème siècle, c'est essentiellement la théorie des fonctions tabulables ; par exemple, résoudre une équation différentielle, c'est trouver sa solution en termes de fonctions élémentaires, ou, sinon, de quadratures. De la même façon, résoudre une équation algébrique, du 16ème au 19ème siècle, c'est essentiellement la résoudre "par radicaux" ; ou sinon, comme le propose Hilbert, à l'aide de fonctions tabulables, ou de ces constructions nomographiques que nous avons oubliées. Au 20ème siècle, résoudre une équation fonctionnelle ou une équation numérique, c'est donner un bon procédé pour approcher rapidement la solution si l'on dispose d'une bonne machine à calculer. Ainsi les notions d'approximation, de rapidité de convergence, de complexité des algorithmes, très liées au développement des moyens de calcul modernes, deviennent-ils beaucoup plus importants. L'ε-entropie de Kolmogorov, par exemple, est une notion de ce type. C'est grâce aux mathématiciens, sans aucun doute, qu'existent les ordinateurs qui ont tant d'importance dans la vie moderne. Mais c'est grâce aux ordinateurs qu'une partie des mathématiques modernes prend tout son sens.

Si cela est vrai, j'ai apporté un élément de réponse à la question posée dans l'introduction. Dans la vie des mathématiques, il y a un facteur de permanence qui est essentiel : il est constitué par les théories acquises, mais plus encore par les problèmes en suspens. Les grands problèmes, à chaque époque, constituent une sorte d'ossature des mathématiques. Mais ce sont des os un peu particuliers, puisque les mathématiques progressent et grandissent en les détruisant. Il y a aussi des facteurs d'évolution, dont les chercheurs n'ont pas nécessairement personnellement conscience, et qui se situent en dehors des mathématiques, dans la vie sociale et en particulier le progrès des techniques. Dans les mathématiques du 20ème siècle, ces facteurs externes semblent être intervenus moins pour créer des sujets, des problèmes, des notions, que pour valoriser une partie des sujets, des problèmes, des concepts dégagés par les mathématiciens. Si nous connaissions mieux l'influence de ces facteurs externes, nous aurions sûrement une meilleure vision de l'histoire des mathématiques, et peut être une meilleure idée de leur développement dans l'avenir.

[1] Proceedings of Symposia in Pure Mathematics, vol. 28. Mathematical developments arising from Hilbert problems. Amer. Math. Soc. 1976.

[2] VITUSHKIN, A. G. On representations of functions by means of superpositions and related topics. Enseign. Math. 23 (1977), 255-320.

[3] KAHANE, J.-P. Sur le treizième problème de Hilbert, le théorème de superposition de Kolmogorov et les sommes algébriques d'arcs croissants, pp. 76-101 in Harmonic Analysis Iraklion 1978, Lecture Notes in Math. 781 (1980).

$X^7 + xX^3 + yX^2 + zX + 1 = 0$

d'Ocagne

$\sqrt{\cdot + \sqrt{\cdot}} + \sqrt{\cdot + \sqrt{\cdot}}$

$P_n(X) = 0$

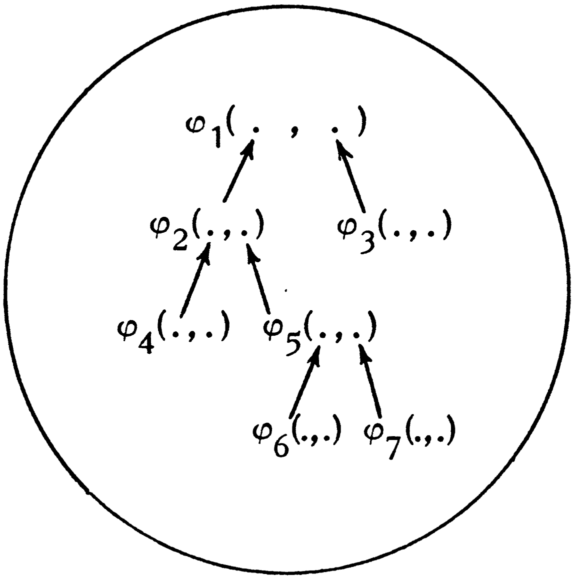
$y = Q_{n-1}(X)$

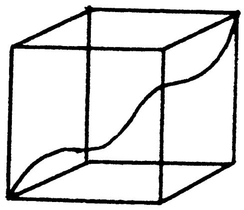
Tschirnhaus

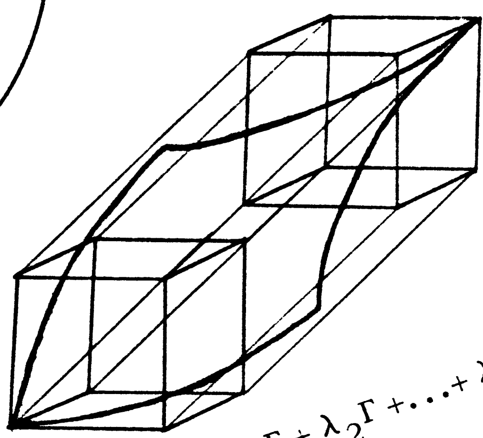
Анатолий Георгиевич Витушкин

Владимир Игоревич Арнольд

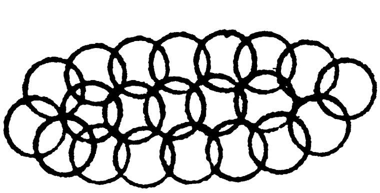
$\Sigma\, c_{pqr}\, x_1^p x_2^q x_3^r$







$H(\varepsilon)$



$(\frac{1}{\varepsilon})^{\frac{n}{\alpha}}$

$E = \lambda_1\Gamma + \lambda_2\Gamma + \ldots + \lambda_n\Gamma$

$f(x_1, \ldots x_n) = \sum_1^{2n+1} g_q\left(\sum_{p=1}^{n} \varphi_{pq}(x_p)\right)$

$C_o(R^n)$.

Андрей Николаевич Колмогоров

NOTES DE LA REDACTION

1 D. HILBERT, *Sur les problèmes futurs des mathématiques*, Compte rendu du deuxième Congrès international des mathématiciens, Paris(Gauthier-Villars), 1902, p.58-114 (p.91-92 : *XIII. Impossibilité de la résolution de l'équation générale du septième degré au moyen de fonctions de deux arguments seulement.*

2 Référence [8] de [2]. Il existe une traduction en anglais dans *Amer. Math. Soc. Translations*, (2), 28(1963), p.51-54.

3 Référence [9] de [2].

4 Publié dans le tome 2(1683) des *Acta Eruditorum*, p.204-207 (voir C.B. Boyer, *A History of Mathematics*, New York(Wiley), 1968, p.472-473, et L.E. Dickson, *Algebraic Theories*, New York(Dover), 1959, p.212).

5 Référence [23] de [2].

6 Référence [4] de [2].

7 Pour ces questions voir C.B. Boyer, *A History of Mathematics*, p.310-317.

8 Le théorème "de Baire" a été énoncé pour $\mathbb{R}$ par W.F. Osgood, en 1897, p.171-173 de *Non-uniform Convergence and the Integration of Series Term by Term* (Amer. Journal Math., 19(1897), 155-190), et, indépendamment, par R. Baire, en 1899, p.65-65 de *Sur les fonctions de variables réelles* (Annali di Mat. pura ed appl., (3), 3(1899), 1-123). Pour $\mathbb{R}^2$ par Osgood, en 1900, p.462 de *Zweite Note über analytische Functionen mehrerer Verändlichen* (Math. Annalen, 53(1900), 461-464), et par Baire, en 1905, pour $\mathbb{R}^n$, p.106 des *Leçons sur les fonctions discontinues*, Paris(Gauthier-Villars).

9 Référence [48] de [2].

10 Référence [37] de [2].

11 Référence [40] de [2].

12 G.G. LORENTZ, *Approximation of Functions*, New York(Holt), 1966, p.168-169.

13 Référence [38] de [2].

14 Ce problème ne figure pas sur la liste des 52 problèmes posés par Luzin à l'époque où il a rédigé sa thèse, problèmes publiés et commentés par N.K. Bari et D.E. Menchov, p.364-386 de Н.Н. Лузин, Интеграл и тригонометрический ряд, Москва(Госуд. издат. тех.-теор. лит.), 1951 .

15 L'article de L. Wantzel, "élève-ingénieur des Ponts-et-chaussées", a paru dans le *Journal des Mathématiques pures et appliquées* de Liouville, t.2(1837), p.366-372.
F. Lindemann, *Ueber die Zahl* π (Math. Annalen, 20(1882), 213-225).

16 Lettres de Cantor des 25 et 29 juin 1877 et de Dedekind du 2 juillet 1877, p.205-216 de la *Philosophie mathématique* de J. Cavaillès, Paris(Hermann), 1962.

17 Référence [52] de [2].