Herbrand et le programme de Hilbert

June 9, 2008

Introduction

Cette note reprend l'exposé que j'ai donné pour la journée sur l''Héritage scientifique de Jacques Herbrand" à l'École Normale Supérieure à l'occasion du centième anniversaire de sa naissance. Cet exposé était, essentiellement, un commentaire de deux preuves de cohérence pour l'arithmétique données dans sa thèse et dans la référence [14]. Ces deux preuves constituent une contribution importante au programme de Hilbert, et je commence par présenter quelles étaient les motivations et les buts de ce programme.

1 La signification des énoncés d'existence en mathématique

Une des meilleures manières de présenter le programme de Hilbert est, comme le fait von Neumann [25], de commencer par une discussion sur la signification des énoncés mathématiques, et en particulier des énoncés qui affirment l'existence d'un objet vérifiant une propriété donnée. Comme le remarque von Neumann, si on a une preuve en mathématique de l'existence d'un nombre réel satisfaisant une propriété $E(x)$, il peut arriver que l'on n'a *aucun* moyen à partir de cette preuve de construire un tel x qui vérifie $E(x)$ (un tel exemple sera donné plus tard). Ce phénomène illustre bien le caractère "non effectif" des mathématiques. On peut noter que cet aspect non effectif des énoncés mathématiques est relativement nouveau : avant 1888 (date du célèbre théorème de la base finie de Hilbert), la plupart des théorèmes d'existence en mathématique contenaient, peut-être implicitement, un algorithme pour construire le témoin (même si le calcul n'était possible peut-être seulement qu'en théorie; Galois, par exemple, insiste sur le caractère purement idéal des calculs nécessaires pour tester si une équation donnée est ou non résoluble par radicaux). On doit noter aussi que cette discussion sur l'effectivité a eu lieu bien avant la définition des fonctions récursives par Church, Kleene et Turing dans les années 1930-40. Comme on peut le voir dans son "Nachlaß"[16], Hilbert était très préoccupé par cette issue dès les années 1890. Une idée d'Hilbert pour analyser ce problème est de remarquer la grande différence entre la "forme" et le "contenu" en mathématique : si la *signification* des énoncés de mathématique n'est peut-être pas si claire et non effective, la *vérification* d'un raisonnement mathématique, elle, doit être parfaitement claire et effective. S'il y a une erreur dans un raisonnement on peut toujours la mettre en évidence, et de plus cette vérification repose uniquement sur la *forme* du raisonnement, et non sur son *contenu*. Il y a donc une différence essentielle en mathématique entre la "sémantique" (qui peut rester vague et intuitive) et la "syntaxe" (qui doit être extrêment précise).

1.1 Exemples

Voyons maintenant quelques exemples en mathématique qui illustrent le problème de la signification des énoncés d'existence. Considérons l'affirmation suivante

Théorème : *Si* $f : [0,1] \to [-1,1]$ *est continue et telle que* $f(0) = -1,\ f(1) = 1$ *alors il existe* x *tel que* $f(x) = 0$ *et* $f(y) < 0$ *si* $0 \leqslant y < x$

On affirme l'existence d'un nombre réel x. La preuve est très courte : il suffit de prendre pour x la borne supérieure de l'ensemble $\{\ y \in [0,1] \mid f(y) < 0\ \}$.

Toutefois, il est impossible d'extraire de cette preuve un moyen de calculer un tel nombre x.

En effet, considérons comme cas particulier la fonction $f : [0,1] \to [-1,1]$ linéaire par morceaux (et clairement calculable) telle que $f(0) = -1,\ f(1) = 1$ et $f(1/3) = -\epsilon, f(2/3) = +\epsilon$ où ϵ est un nombre positif ou nul très près de 0. On voit alors que le nombre x doit être proche de $1/3$ ou proche de $1/2$ suivant que ϵ est positif ou nul. Mais si on sait rien de plus sur ϵ, on ne peut pas décider.

Par exemple, prenons $\epsilon = \Sigma_n \frac{\epsilon_n}{2^n}$ avec $\epsilon_k = 0$ si $2k$ est une somme de deux nombres premiers et $\epsilon_k = 1$ dans le cas contraire. Alors ϵ est calculable : on peut en calculer des approximations arbitrairement proches. Mais on ne peut trouver x tel que $f(x) = 0$ et $f(y) < 0$ si $y < x$. Il est même impossible une approximation de x à $1/12$ près à partir de la preuve d"'existence" d'un tel objet. En effet, trouver une telle approximation revient à décider la conjecture de Goldbach, et il est clair que la preuve très courte d'existence d'un nombre x que l'on a donné ne contient aucune indication sur ce problème!

Cet exemple peut paraitre artificiel, mais il illustre bien la subtilité de la notion d'existence en mathématique. Voici d'autres exemples plus naturels.

En algèbre (théorème de la base finie) Hilbert utilisait dans le cas de base le fait suivant : si on a une suite d'entiers $n_1,\ n_2,\ \ldots$ il existe k tel que n_k soit minimum (on a $n_l \geqslant n_k$ pour tout l). Peut-on calculer un tel indice k ?

En théorie des nombres Dirichlet a une preuve en utilisant de l'analyse qu'il y a un nombre infini de nombres premiers de la forme $an + b,\ a,\ b$ premiers entre eux. Peut-on extraire de cette preuve *un* tel nombre premier (étant donné a et b) ?

En algèbre réelle, Artin et Schreier présentaient une solution du XVIIème problème de Hilbert

Théorème : *Un polynome* P *de* $\mathbb{Q}[X_1, \ldots, X_n]$ *qui ne prend que des valeurs* $\geqslant 0$ *peut s'écrire comme somme de carrés* de fractions rationnelles[1].

La preuve utilise des raisonnements transfinis (existence d'un bon ordre sur tout ensemble). Une question naturelle (qui était posée par Artin) est alors la suivante : si on donne le polynome P peut-on l'écrire effectivement comme une somme de carrés à partir de cette preuve ?

Terminons par un exemple en algèbre commutative : le théorème de Quillen-Suslin (problème de Serre).

Théorème : *Une matrice idempotente de polynomes est semblable à une matrice de projection canonique.*

La preuve de Suslin utilise l'existence d'un idéal maximal. La preuve de Quillen utilise des résultats de nature "locale-globale", qui reposent sur l'existence d'idéaux premiers. Est-il possible d'extraire de ces preuves un algorithme pour résoudre la question suivante[2] : si on a une matrice de polynomes explicitement donnée M, peut-on calculer effectivement P tel que PMP^{-1} soit une matrice de projection canonique ?

[1] En général, il peut ne pas être une somme de carrés de *polynomes*.

[2] Pour une analyse de ces questions, voir [22] et [30].

1.2 Quelques questions

Ces exemples suggèrent les questions suivantes. Tout d'abord, comme on l'a vu par le premier exemple, on ne peut extraire en général de manière effective un témoin à partir d'une preuve d'existence en mathématique. Y-a-t'il certains cas où on le peut ? Ensuite, quelle est la signification intuitive concrète[3] de l'énoncé $\exists x.E(x)$, qui doit être plus subtile en général que la seule existence de x ?

2 Critique de Brouwer et le programme de Hilbert

Comme Hilbert, Brouwer prend très au sérieux ces problèmes de signification des énoncés mathématiques. Le résultat de son analyse est qu'il faut reconstruire les mathématiques suivant des critères plus rigoureux. Il reconnait la source du caractère non effectif des mathématiques dans le principe du tiers-exclu[4]. Ce principe permet de montrer l'existence d'un élément vérifiant une propriété en montrant qu'il est impossible que tout élément vérifie sa négation. Il propose donc de rejeter ce principe. Cette restriction des mathématiques aux raisonnements qui n'utilisent pas ce principe (les mathématiques "finitistes", "constructives" ou "intuitionistes"; comme Herbrand, nous utiliserons ces termes de manière équivalente) est absolument impensable pour Hilbert. Il en fait la critique suivante [18] (mes italiques) : "*Si nous restons dans le domaine des propositions finitistes,* comme nous le devons d'ailleurs*, les relations logiques qui y règnent manquent singulièrement de perspicuité, et ce défaut s'aggrave au point de devenir insupportable lorsque "tous" et "il existe" se combinent . . . il est de fait que personne, même qui parlerait le dialecte des anges, n'empêchera les hommes de nier des assertions quelconques . . . et d'appliquer le tiers exclu. Que faire ?*". On voit dans cette citation qu'Hilbert reconnait tout à fait le problème de la signification des énoncés mathématiques. Si on veut des énoncés qui ont un sens clair, il faut se restreindre aux énoncés finitistes. Le problème pour Hilbert est alors que ces énoncés, et encore plus les preuves, deviennent complexes au point d'être inutilisables.

La solution de Hilbert est très originale. On veut garder la forme usuelle des énoncés et preuves mathématiques et on a vu que cet aspect formel est très précis, peut se décrire de manière constructive, et suffit à lui-même pour vérifier la correction des arguments mathématiques. On va donc simplement oublier l'aspect sémantique, ou si l'on veut, remplacer la "sémantique" par la "syntaxe". Le calcul logique, qui est bien défini et gouverne les mathématiques classiques, est considéré comme un jeu formel et sans contenu.

On notera l'analogie avec l'analyse de la théorie physique, qui avait été dégagée peu auparavant par Duhem [10]. Comme le présente Duhem, on met au point un modèle purement formel de certains phénomènes physiques et seul compte la condition que les observations prévues par ce modèle se vérifient. Le fait que les résultats intermédiaires du modèle, peut-être non observables, peuvent ne correspondre à aucune propriété du monde physique n'a aucune importance. Ce qui est joue un rôle essentiel par contre sont les critères d'élégance de la théorie et la facilité avec laquelle on peut prévoir des résultats observables. La situation est similaire ici : si le critère d'élégance est crucial (et motive le programme de Hilbert) il est en outre essentiel que si l'on prouve de manière formelle un énoncé finitiste, que l'on peut vérifier, alors cet énoncé est effectivement correct. Comme le remarque Hilbert, ceci revient à prouver la *cohérence* de ce calcul formel. On saura alors plus généralement que tous les énoncés purement universels que

[3] Cette question a été analysée par Gentzen, 1936, pour les énoncés d'arithmétique.

[4] Cette analyse, en 1908, n'était pas triviale dans le contexte de l'époque, car la discussion se concentrait alors autour de l'axiome du choix et de la preuve de Zermelo de l'existence d'un bon ordre sur tout ensemble à partir de cet axiome. Les discussions semblent attribuer cette existence non effective à l'utilisation de l'axiome du choix, et non à l'utilisation du tiers-exclu. Pour une discussion de ce point, voir l'introduction du livre de Bishop [5].

l'on prouve de manière formelle, peut-être en utilisant des arguments transfinis, sont corrects[5] : en effet, si le calcul est cohérent, on ne peut pas avoir de contre-exemples. On peut remarquer que cet énoncé de cohérence est lui-même un énoncé purement universel. Si on arrive de plus à montrer cette cohérence en n'utilisant que des raisonnements intuitionnistes, on aura vraiment répondu à la critique de Brouwer. Tel est, en gros, le programme de Hilbert, auquel Herbrand va apporter des contributions essentielles.

2.1 Formalisation des mathématiques

La première étape consiste à préciser complètement la structure logique des raisonnements mathématiques. Une très bonne description de la situation avant les travaux d'Herbrand est donnée par Weyl [31]. Il présente une stratification des raisonnements en deux calculs. Le premier calcul, le calcul *propositionnel*, ou calcul Booléen, est décidable (au moins en théorie) par la méthode des tables de vérité. Cette méthode donne à la fois une *sémantique* et un *procédé de décision*. Le deuxième calcul, le calcul *des prédicats*, est obtenu en ajoutant les quantifications existentielles et universelles. En général on ne peut pas calculer la valeur de vérité d'un énoncé pour ce calcul et, suivant la critique de Brouwer, les énoncés avec quantificateurs (sur un domaine qui peut être infini) n'ont pas un sens clair. On doit se restreindre à une présentation syntaxique précise des règles de déduction qui permettent de manipuler ces énoncés de manière formelle. On pourra considérer que les énoncés sans quantificateurs correspondent aux énoncés "finitistes".

Hilbert a en fait aussi une approche originale du calcul des prédicats. Elle consiste à définir les quantificateurs à partir du symbole τ qui a pour unique axiome $A(\tau_x A) \rightarrow A(x)$ ce qui constitue une forme de l'axiome du choix[6]. Ce symbole et cet axiome concentrent le coté non effectif des quantifications : même si A est décidable, il n'y a en général aucun moyen de trouver une valeur de $\tau_x A$. Le but est de montrer que l'on peut éliminer ce symbole dans toute démonstration donnée d'un résultat finitaire[7].

2.2 Représentation d'un ensemble infini

Le premier problème considéré par Hilbert [17] est celui de la théorie la plus simple d'un ensemble infini, qui est donné dans le langage avec un symbole de constante 0, un symbole de fonction $S(x)$ et les deux axiomes

$$S(x) \neq 0, \qquad S(x) = S(y) \rightarrow x = y$$

qui expriment que cette fonction n'est pas surjective mais est injective. Il n'y a pas de modèle fini de cette théorie. "Naivement", ces axiomes sont vérifiés pour $\mathbb{N}$; il y a donc intuitivement un modèle (infini), ce qui montrerait la cohérence. Mais cet argument est sans valeur pour un finitiste, car la signification des quantificateurs est problématique, et on ne peut pas calculer a priori la valeur de vérité des énoncés quantifiés. Il n'est donc pas du tout clair pour un finiste que les deux axiomes ne conduisent pas à une contradiction si on les utilise dans le calcul des prédicats. Pour être vraiment écrit dans le calcul des prédicats, la théorie doit être complémentée par les axiomes suivants qui axiomatisent l'égalité.

$$x = x, \quad x = y \wedge y = z \rightarrow x = z, \quad x = y \rightarrow y = x$$

$$x = y \rightarrow S(x) = S(y)$$

[5] Hilbert donne comme exemple de tel énoncé le Théorème de Fermat.

[6] On peut alors définir $\forall x.A$ par $A(\tau_x A)$.

[7] Le symbole τ sera remplacé par la suite par un symbole dual ϵ, qui dénote une fonction de choix. Bourbaki adoptera le symbole τ dans sa formulation de la théorie des ensembles, mais avec la signification duale du symbole de choix ϵ. La méthode de Hilbert est expliquée de manière très suggestive dans la référence [32].

2.3 Premières preuves de cohérence

L'argument de Hilbert (1904) est en gros le suivant. On regarde la théorie précédente de manière purement syntaxique. Les termes $t, u, \ldots$ sont 0, $S(0)$, $S(S(0)), \ldots$ et les formules que l'on considère des équations $t = u$. On voit que les axiomes $a = a$ sont des équations vraies. Les autres axiomes peuvent être vus comme des règles d'inférences qui permettent de déduire d'autres équations à partir d'équations. Par exemple, l'axiome $x = y \to y = x$ s'interprète de la manière suivante : si on a déduit $t = u$ alors on peut ajouter aussi $u = t$ dans les conséquences possibles. On voit directement que l'on ne pourra jamais déduire que des équations vraies : si par exemple on utilise la règle

$$x = y \wedge y = z \to x = z$$

en déduisant $t = v$ à partir de $t = u$ et de $u = v$, il est clair que si $t = u$ est vrai (c'est-à-dire t et u sont le même terme $S^k(0)$) et $u = v$ est vrai alors $t = v$ est aussi vrai. Par ce moyen il semble que l'on puisse montrer la cohérence d'une théorie qui n'a pas de modèle fini, à partir de raisonnement purement syntaxique. (Dans ce raisonnement, on ne considère jamais l'ensemble des termes comme une totalité "close".)

Hilbert était très optimiste sur la portée d'une telle méthode syntaxique. Il conclut son papier [17] : "*De la même manière, nous pouvons montrer que les notions fondamentales de la théorie de Cantor, en particulier les alephs, ont une existence cohérente.*"

Le problème dans cette approche est que cet argument ne considère qu'une partie des raisonnements possibles en calcul des prédicats, ceux qui utilisent directement les axiomes comme règle d'inférence pour conclure des égalités entre des termes clos. Mais on peut avoir des raisonnements indirects, en passant par des lemmes qui utilisent des énoncés quantifiés complexes. Comme on l'a vu, le sens de ces énoncés quantifiés n'est pas si clair et il se pourrait bien a priori, qu'en utilisant un tel énoncé A et les règles de déduction du calcul des prédicat, on arrive à montrer A à partir des axiomes et aussi on arrive à déduire une contradiction à partir de A.

Une preuve de cohérence finitiste pour le calcul des prédicats étendus avec les axiomes de la section 2.2 est obtenue par von Neumann [24], qui utilise l'approche de Hilbert du calcul des prédicat (avec le symbole τ) et qui précise un essai précédent du à Ackermann [1, 33].

3 Première contribution d'Herbrand

Le travail de thèse d'Herbrand contient une preuve remarquablement simple de cohérence de la théorie précédente. Herbrand montre en fait la cohérence d'une *extension* de cette théorie où l'on ajoute les axiomes d'induction

$$\forall \vec{x}.\ \phi(\vec{x}, 0) \wedge (\forall y.\ \phi(\vec{x}, y) \to \phi(\vec{x}, S(y))) \to \forall y. \phi(\vec{x}, y)$$

L'argument est de plus d'une simplicité remarquable, comparée aux preuves d'Ackermann et von Neumann. Il fournit non seulement une preuve de cohérence mais aussi une *caractérisation complète* de la théorie (procédure de décision). L'approche consiste à "éliminer les quantificateurs" en associant à chaque formule $\phi(\vec{x})$ une autre formule $\phi'(\vec{x})$ *sans quantificateurs* telle que

$$\forall \vec{x}.\ \phi(\vec{x}) \leftrightarrow \phi'(\vec{x})$$

est prouvable à partir des axiomes donnés. Un corollaire est que, dans ce cas particulier, on peut en fait calculer la valeur de vérité des énoncés quantifiés, bien que la quantification porte sur un domaine infini.

Herbrand a le commentaire suivant : "*Nous allons faire sur ce terme les opérations qui, en algèbre ordinaire, correspondent à l'élimination de x dans un système dégalités et d'inégalités.*"

Il ajoute aussi, faisant allusion à la théorie récente d'Artin et Schreier : "*La méthode employée dans ce chapitre est susceptible d'autres applications ; elle fournit toujours, en même temps que la non-contradiction de la théorie étudiée, sa résolubilité ... Il nous parait probable qu'elle permettrait également d'arriver à la non-contradiction de la théorie des corps réels et "réellement fermés"; mais les méthodes du Chapitre suivant nous y conduiraient plus aisément.*" C'est une formulation du résultat d'élimination des quantificateurs sur la théorie des corps réels clos, qui sera publié par Tarski quelques années plus tard[8]. Citons le travail [8], pour une analyse récente de la situation, et par lequel on peut comprendre la remarque d'Herbrand, que l'autre méthode qu'il va développer (et que l'on va décrire dans la prochaine section), permet une approche alternative pour montrer la cohérence de la théorie des corps réels clos.

4 Deuxième preuve de cohérence et Thèorème Fondamental

Le problème avec cette première approche est qu'elle ne peut pas marcher dès que l'on ajoute l'addition et la multiplication. En effet, si elle s'appliquait, on aurait aussi un procédé de décision pour cette extension. Mais il est intuitif (et ce sera précisé plus tard par Church [6]) qu'un tel procédé de décision ne peut pas exister, puisque l'on peut par exemple exprimer dans cette théorie le théorème de Fermat pour un exposant donné. (Il faut noter aussi que par contre, si on se restreint à ajouter seulement l'addition, la méthode d'élimination des quantificateurs marche bien, comme l'a montré Pressburger.)

La deuxième méthode proposée par Herbrand résout ces difficultés. Cette méthode est extrèmement flexible et marche généralement en ajoutant des symboles de fonctions avec des axiomes qui spécifient (de manière intuitionniste) ces fonctions. Par exemple, elle marchera pour la théorie considérée en section 2.2 étendue avec les axiomes

$$x + 0 = x, \quad x + S(y) = S(x + y), \quad x = z \wedge y = t \rightarrow x + y = z + t$$

$$x \times 0 = 0, \quad x \times S(y) = x + x \times y, , \quad x = z \wedge y = t \rightarrow x \times y = z \times t$$

Dans une telle théorie, On peut montrer que si $\psi(x)$ est sans quantificateur et t est un terme clos, alors soit $\psi(t)$, soit $\neg\ \psi(t)$ est prouvable (intuitivement, les formules sans quantificateurs sont décidables). Notons que pour cette théorie utilisée par Herbrand, tout terme clos est prouvablement égal à un terme de la forme $0,\ S(0),\ S(S(0)), \ldots$ Les substitutions closes à effectuer ont donc une structure très simple.

On peut ajouter des définitions de la forme

$$\psi(x) \rightarrow f(x) = 0, \quad (\neg\psi(x)) \rightarrow f(x) = S(0), \quad x = y \rightarrow f(x) = f(y)$$

pour $\psi(x)$ formule sans quantificateurs. En effet, ceci est une spécification complète d'une fonction que l'on peut calculer, puisque la propriété $\psi(t)$ est décidable. Il n'y a pas d'axiome d'induction explicite comme dans la version précédente, mais on peut montrer aussi que l'axiome d'induction est prouvable pour les énoncés sans quantificateurs en utilisant de telles fonctions.

La théorie de l'arithmétique considérée par Herbrand est une théorie *universelle*, c'est-à-dire n'ayant que des axiomes de la forme $\forall\vec{x}.\phi(\vec{x})$ où ϕ est sans quantificateurs.

Ce qui est remarquable, c'est que cette méthode marchera sans avoir à décider tous les énoncés. Elle repose sur le Théorème Fondamental suivant[9].

[8] Il semble que Tarski ait obtenu ce résulat de manière indépendent des travaux d'Artin et Schreier.

[9] La formulation que je donne est en fait un cas particulier du résultat d'Herbrand; on suppose que l'on a au mois un symbole de constante. Une preuve claire de ce résultat est présentée dans le livre de Shoenfield [27].

Théorème : *Une théorie purement universelle, c'est-à-dire n'ayant que des axiomes de la forme* $\forall \vec{x}.\phi(\vec{x})$ *où* ϕ *est sans quantificateurs, est* cohérente *si, et seulement si, la théorie* propositionnelle *qui est formée par tous les substitutions closes* $\phi(\vec{t})$ *est cohérente.*

Ce résultat réduit, en un certain sens, le calcul des prédicats (non finitiste) au calcul propositionnel (finitiste). Ceci joue un rôle essentiel en démonstration automatique. On voit que ce théorème est précisément ce qui manque à l'esquisse de la preuve de Hilbert 1904 pour être conclusive.

Ce théorème ne fait intervenir *que des notions syntaxiques* : dérivation dans la théorie des énoncés avec quantificateurs et dérivation en calcul propositionnel. L'énoncé suit donc bien les restrictions intuitionistes nécessaires, contrairement aux travaux précédents de Lowenheim et Skolem. La preuve de ce théorème n'est pas du tout évidente : on doit partir d'une dérivation en calcul avec quantificateurs et la transformer progressivement en une dérivation sans quantificateurs. En fait, la preuve qu'Herbrand en donne est *fausse*[10] !

4.1 Théorème fondamental, exemple

Voici un exemple simple qui illustre le Théorème Fondamental. Considérons la théorie :

$$\forall x\ y\ z.\ \ x \leqslant y \wedge y \leqslant z \ \ \rightarrow \ \ x \leqslant z, \qquad \forall x.\ \ x \leqslant x, \ \ \forall x\ y.\ \ x \leqslant f(x,y), \quad \forall x\ y.\ \ y \leqslant f(x,y)$$

$$\forall x.\ \ \neg(a \leqslant x \wedge b \leqslant x \wedge c \leqslant x)$$

C'est une théorie contradictoire. De manière sémantique, si on a un préordre et un majorant pour deux éléments, on a un majorant pour trois éléments et donc le dernier axiome est incompatible avec les précédent. Le Théorème Fondamental dit que l'on doit observer cette contradiction à un niveau purement propositionnel en regardant les instantiations closes de ces axiomes. En effet parmi ces instantiations, nous avons

$$a \leqslant f(a,b), \quad b \leqslant f(a,b), \quad f(a,b) \leqslant f(f(a,b),c), \quad c \leqslant f(f(a,b),c)$$

$$a \leqslant f(a,b) \wedge f(a,b) \leqslant f(f(a,b),c) \rightarrow a \leqslant f(f(a,b),c)$$

$$b \leqslant f(a,b) \wedge f(a,b) \leqslant f(f(a,b),c) \rightarrow b \leqslant f(f(a,b),c)$$

$$\neg(a \leqslant f(f(a,b),c) \wedge \ b \leqslant f(f(a,b),c) \wedge \ c \leqslant f(f(a,b),c))$$

qui sont contradictoires de manière purement propositionelle (sans faire intervenir d'énoncés quantifiés).

L'application du Théorème fondamental à l'arithmétique est directe. La théorie ne comprend bien que des axiomes purement universels. De plus, si on considère les instantiations de ces axiomes, on n'obtient que des énoncés vrais (exactement comme dans l'esquisse proposée par Hilbert [17]), et donc on ne peut pas obtenir de contradiction au niveau propositionel.

[10] Une très bonne discussion sur ce point est accessible à la page web de Peter Andrews. Voir aussi le papier de Tait [29]. En gros, l'argument d'Herbrand ne prévoit pas de différence de tailles entre la preuve de contradiction en calcul des prédicats et en calcul propositionel, alors que l'on peut donner des exemples où la preuve grossit de manière non élémentaire (tour d'exponentielle). Ce phénomène illustre bien le gain obtenu en utilisant les moyens idéaux donnés par les quantificateurs. Comme il m'a été signalé par R. Zach, la première preuve finitiste correcte du résultat d'Herbrand semble être due à Bernays [4], qui utilise la technique présentée dans [24].

4.2 Un corollaire remarquable

Le résultat suivant est un corollaire du Théorème Fondamental qui sera explicité par Bernays [4].

Corollaire : *Si on montre $\exists x.\psi(x)$ dans l'arithmétique présentée par Herbrand et $\psi(x)$ est sans quantificateur, alors on peut calculer t tel que $\psi(t)$ est prouvable*

En effet, la théorie universelle en ajoutant l'axiome $\forall x.\neg\psi(x)$ est contradictoire. Par le Théorème Fondamental, on a donc (explicitement) une contradiction dans la théorie propositionelle où l'on a ajouté toutes les instantiations $\neg\psi(0),\ \neg\psi(S(0)),\ldots$ Cette contradiction explicite est un objet fini qui fournit un témoin t.

Ce résulat remarquable est utilisé par Church [6] pour une présentation finitiste du fait qu'il n'y a pas d'algorithme de décision pour le calcul des prédicats. L'idée de cette application est la suivante. On peut former par exemple l'énoncé $A = \exists a\ b\ c.\ S(a)^3 + S(b)^3 = S(c)^3$. En utilisant le Corollaire du Théorème fondamental, on voit que cet énoncé est prouvable dans l'arithmétique d'Herbrand si, et seulement si, on peut trouver $p, q, r > 0$ tels que $p^3 + q^3 = r^3$. Un algorithme de décision pour le calcul des prédicats donnerait donc un moyen uniforme de décider tous les énoncés existentiels en arithmétique, ce que Church avait montré impossible auparavant.

Ce Corollaire suggère le principe (heuristique) général suivant.

Principe : *Si on montre de manière classique l'existence d'un objet "*concret*" (entier, ou qui peut se coder comme un entier) qui vérifie une propriété* décidable *alors on peut extraire de cette preuve un moyen de calculer cet objet.*

Ceci n'est *pas* vérifié pour les énoncés de la forme $\exists x.\forall y.\phi(x,y)$. Par exemple, il est très simple de montrer $\exists x.\forall y.g(x) \leqslant g(y)$, mais il est impossible à partir de cette preuve de calculer un témoin comme fonction de g.

4.3 Généralisation

Herbrand énonce son résulat sans restriction sur la forme des axiomes de la théorie. Il est possible en effet de se ramener au cas d'une théorie purement universelle en introduisant des symboles de fonctions. Prenons l'exemple suivant : supposons que l'on a une preuve en calcul des prédicats de la formule $\exists u.\forall t.G(u,t)$ à partir de l'axiome unique $\forall x.\exists y.\forall z.\exists w.F(x,y,z,w)$ où les formules $G(u,t)$ et $F(x,y,z,w)$ sont sans quantificateurs. Ceci signifie que la théorie avec les deux axiomes

$$\forall x.\exists y.\forall z.\exists w.F(x,y,z,w),\quad \forall u.\exists t.\neg\ G(u,t)$$

est contradictoire. Cette théorie se transforme en une théorie universelle en introduisant des symboles de fonction $f_1(x),\ f_2(x,z),\ f_3(u)$, et on a que la théorie

$$\forall xz.F(x,f_1(x),z,f_2(x,z)),\quad \forall u.\neg\ G(u,f_3(u))$$

est aussi contradictoire. D'après le Thérorème Fondamental, ceci revient à dire que la théorie propositionnelle obtenue en remplaçant x, z, u par les termes générés par l'application répétée des fonctions $f_1,\ f_2,\ f_3$ sur une constantes intiale a

$$a,\ f_1(a),\ f_2(a,a), f_3(a),\ \ldots,\ f_2(f_1(a),f_2(a,a)),\ldots$$

est contradictoire. Comme le remarque A. Robinson [23], ceci constitue un moyen possible pour chercher les preuves en calcul des prédicats de manière automatique. De plus, cette approche est en fait assez proche de la manière dont procèdent les mathématiciens par exemple

en géométrie élémentaire : les termes obtenus par les fonctions $f_1(x),\ f_2(x,z),\ f_3(u)$ correspondent aux points, lignes ... auxiliaires construits pour résoudre un problème. Cette approche est absolument fondamentale en démonstration automatique, le problème principal étant de réduire autant que possible la production des termes auxiliaires[11].

5 Qu'est-ce qu'un algorithme ?

Le travail d'Herbrand [15] va jouer aussi un rôle essentiel dans la définition mathématique de la notion d'algorithme qui va être mise au point par Church, Turing, Gödel, Kleene, Post dans les années 1930-40. Herbrand présente l'arithmétique comme un système "ouvert" : on peut rajouter de nouvelles fonctions avec leur spécifications. La seule condition est que cette spécification permette le calcul de manière intuitioniste de la fonction. Herbrand insiste de plus sur le fait que ce caractère ouvert du système est inévitable : il est *impossible* de donner un moyen uniforme de décrire toutes les fonctions calculables. C'est un argument très simple de diagonalisation : si l'on a une énumération intuitioniste f_n de fonctions calculables, alors la fonction $n \longmapsto 1 + f_n(n)$ est aussi calculable, mais ne figure pas dans cette énumération. Ceci explique, d'après Herbrand, comment on échappe au résultat de Gödel sur la non existence d'une preuve de cohérence : elle ne s'applique que pour un système fixe; mais si l'on fixe le système on peut effectuer la preuve de cohérence dans un système étendu où l'on a ajouté de nouvelles fonctions. Ces reflexions contribueront à la mise au point d'une définition générale de la notion d'algorithme (par Church, Kleene et Turing). En réflechissant á cette introduction de fonctions intuitionistes, Gödel mettra au point quelques années plus tard la notion de fonctions d'"Herbrand-Gödel", qui est une définition possible des fonctions récursives[12].

Conclusion

Les deux travaux d'Herbrand que nous avons décrits constituent une contribution essentielle au programme de Hilbert. Ils ont eu une importance cruciale en théorie de la récursivité et en démonstration automatique. Peu après ces travaux, Kolmogorov et Heyting montrent comment formaliser la logique intuitioniste (un résultat tout à fait inattendu), et on peut alors analyser le calcul des prédicats comme le calcul des prédicats intuitionistes auquel on ajoute le principe du tiers-exclu. À l'aide de cette formalisation, et en s'appuyant aussi sur la présentation d'Herbrand, Gödel donne un moyen très simple (la traduction négative [12]) qui permet d'interpréter l'arithmétique non finitiste en arithmétique intuitionniste. Il est alors naturel d'essayer d'étendre cette interprétation à l'Analyse, mais une telle interprétation ne marche pas pour l'axiome du choix dépendent. Ce problème sera résolu par Spector [28] (voir [3] et [19] pour une autre interprétation possible). Plus récemment, J.L. Krivine a annoncé une interprétation calculatoire possible pour le système ZFC, qui étend son interprétation précédente [20], à l'axiome du choix général. Mais même la compréhension du calcul des prédicats réserve encore des surprises, comme le montre par exemple le travail [21], qui analyse la signification de tautologie classique du genre $\exists x.\forall y.\ (P(x) \to P(y))$.

[11] Il faut noter que l'idée de cette approche est essentiellement contenue dans les travaux de Lowenheim et Skolem, qui sont antérieurs aux travaux d'Herbrand, cf. l'introduction de [26]. Toutefois, Lowenheim et Skolem utilisent des notions qui ne sont pas interprétables de manière finitiste.

[12] Gödel apporte quelques modifications significatives à l'approche d'Herbrand : la spécification de la fonction doit être faite par un système d' équations uniquement, la preuve du fait que ce système défini bien une fonction peut être non finitiste, et enfin, le calcul d'une valeur de cette fonction peut se faire par reécriture uniquement à partir du système d'équations.

References

[1] Ackermann, W. Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit. *Math. Ann.* 93, pp. 1–36, 1924.

[2] P. Benacerraf and H. Putnam. Philosophy of mathematics. Selected readings. Cambridge University Press, Second Edition, 1985.

[3] Berardi, S. et Bezem, M. et Coquand, T. On the computational content of the axiom of choice. *J. Symbolic Logic 63* (1998), no. 2, 600–622.

[4] P. Bernays. Lectures at Princeton. Mimeographed notes, The Institute for Advanced Study, N.J, 1936.

[5] E. Bishop. *Foundations of constructive analysis.* McGraw-Hill Book Co., New York-Toronto, Ont.-London 1967.

[6] A. Church. A Note on the Entscheidungsproblem. *The Journal of Symbolic Logic*, Volume 1, Number 1, 1936.

[7] A. Church. Correction to *A Note on the Entscheidungsproblem. The Journal of Symbolic Logic*, Volume 1, Number 2, 1936.

[8] M. Coste, H. Lombardi et M.F. Roy. Dynamical method in algebra: effective Nullstellesätze. *Annals of Pure and Applied Logic*, 111, 203-256 (2001).

[9] J. Dubucs et P. Égré. Jacques Herbrand. in M. Bithold et J. Gayon, dir., *L'épistémologie française, 1850-1950*, Paris, PUF, à paraitre.

[10] P. Duhem. *La théorie physique : son objet et sa structure.* Paris, Chevalier et Rivière, 1906.

[11] Gentzen, G. The collected papers of Gerhard Gentzen. Edited by M. E. Szabo.North-Holland 1969.

[12] Gödel, K. Zur intuitionistichen Aritmetik und Zahlentheorie. *Ergebnisseneis mathematischen Kolloquiums* 4. 34–38, 1933.

[13] van Heijenoort, J. *From Frege to Gödel. A source book in mathematical logic, 1879–1931.* Reprint of the third printing of the 1967 original. Edited by Jean van Heijenoort. Harvard University Press, Cambridge, MA, 2002.

[14] J. Herbrand. *Recherche sur la théorie de la démonstration.* Thèse à l'Université de Paris, 1930.

[15] J. Herbrand. Sur la non-contradiction de l'arithmétique. *Journal für die reine und angewandte Mathematik 166*, 1-8 (1931).

[16] David Hilbert's Mathematical Notebooks. Maintenu par S. Hayashi et accessible à wwww.shayashi.jb/HNBP/index.html.

[17] D. Hilbert. Über die Grundlagen der Logik und der Arithmetik. In Verhandlungen des dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904. English translation in van Heijenoort. ??

[18] D. Hilbert. *Sur l'infini* Math. Annal. 95, 1926, p. 161-190. Traduction francaise dans Largeault.

[19] J.L. Krivine. Dependent choice, 'quote' and the clock. Th. Comp. Sc., 308, p. 259-276 (2003).

[20] J.L. Krivine. Typed lambda-calculus in classical Zermelo-Fraenkel set theory. *Arch. Math. Log.*, 40, 3, p. 189-205 (2001).

[21] J.L. Krivine et Y. Legrandgérard. Formules valides, jeux et protocoles réseaux. A paraitre, 2008.

[22] H. Lombardi, S. Barhoumi et I. Yengui. Projective modules over polynomial rings: a constructive approach. to appear in Math. Nachrichten, 2008.

[23] A. Robinson. Proving Theorems (as Done by Man, Logician, or Machine). Summaries of Talks Presented at the Summer Institute for Symbolic Logic, Princeton, New Jersey, 1957.

[24] von Neumann, Zur Hilbertschen Beweistheorie *Math.Z.* 26, pp. 1–46

[25] J. von Neumann. Die formalistische Grundlegung der Mathematik. *Erkenntnis*, vol. 2. Translated in English in Benacerraf and Putnam.

[26] *Automation of Reasoning, Classical Papers on Computation Logic 1, 1957-1966.* Edited by J. Siekman and G. Wrightson, Springer-Verlag, 1983.

[27] J.R. Shoenfield. *Mathematical Logic.* Addison-Wesley, 1967.

[28] Spector, C. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In *Proc. Sympos. Pure Math.* Vol. V pp. 1–27 American Mathematical Society, Providence, R.I, 1962.

[29] . W.Tait. Gödel's Correspondence on Proof Theory and Constructive Mathematics. *Philosophia Mathematica.* 2006; 14: 76-111

[30] I. Yengui. Making the use of maximal ideals constructive. *Theor. Comp. Sci.* 293(1-3): 174-178 (2008).

[31] H. Weyl. Consistency in Mathematics. *Rice Institut Pamphlet* 16, p. 245-265, 1929.

[32] Weyl, H. David Hilbert and his mathematical work. *Bull. Amer. Math. Soc.* 50, (1944). 612–654.

[33] Zach, R. Numbers and functions in Hilbert's finitism. *Taiwanese J. Philos. Hist. Sci.* 10 (1998), 33–60.