

### 5.3. Intuitionnistes contre formalistes

#### 5.3.1. Les intuitionnistes

Au lieu de tenter de résoudre les problèmes des mathématiques par des constructions formelles où l'intuition n'aurait plus sa place, les intuitionnistes souhaitent revenir en arrière et ne conserver des mathématiques que la partie intuitivement accessible. Ils refusaient en particulier la notion d'infini achevé : pour eux, l'existence d'un ensemble achevé contenant tous les nombres naturels n'était pas évidente. Seul existait le fait qu'on pouvait toujours construire un nombre plus grand que n'importe quel autre nombre. Tous les objets mathématiques devaient se construire<sup>1</sup>.

Ce type d'attitude conduisit Brouwer<sup>2</sup> à contester certains principes considérés jusque-là comme évidents, comme le principe du tiers exclus<sup>3</sup>. Brouwer reconnaissait volontiers que dans le cas où le domaine  $\mathcal{D}$  que parcourt un prédicat  $P(x)$  est fini,  $\exists xP(x) \vee \neg \exists xP(x)$  est vrai. Il est en effet possible dans ce cas de vérifier pour chaque élément de  $\mathcal{D}$  si l'on a  $P(x)$ .

En revanche, pour un ensemble  $\mathcal{D}$  infini,  $\exists xP(x) \vee \neg \exists xP(x)$  n'est pas forcément vrai. Il nous est en effet matériellement impossible de vérifier pour tous les éléments de  $\mathcal{D}$  si l'on a  $P(x)$ . Bien sûr, nous pouvons, avec de la chance, trouver un tel  $x$  rapidement, mais nous pouvons aussi bien ne jamais le trouver.

Il y a une autre conséquence des notions intuitionnistes : pour démontrer  $\exists xP(x)$ , il faut trouver un tel  $x$  et l'exhiber. Une démonstration par l'absurde, qui montrerait la fausseté de  $\neg \exists xP(x)$  ne montrerait en aucun cas qu'un tel  $x$  existe mais seulement que  $\neg \neg \exists xP(x)$  est vraie.

On voit donc que pour un intuitionniste, dans un ensemble infini,  $\neg \neg A$  n'est pas systématiquement équivalent à  $A$ .

Depuis 1918, les mathématiques intuitionnistes se sont développées parallèlement aux mathématiques "traditionnelles". Les démonstrations constructivistes donnent souvent des résultats plus tangibles que les démonstrations formelles, mais sont plus difficiles. Il faut reconnaître que les intuitionnistes sont restés en deçà des résultats formalistes.

Leur apport fut néanmoins très important, surtout sur le plan épistémologique, car il conduisit à une réflexion profonde sur la façon d'appréhender l'infini.

#### 5.3.2. Les formalistes

L'école formaliste emmenée par le mathématicien David Hilbert choisit une voie totalement différente de la voie intuitionniste.

Constatant que les paradoxes venaient de systèmes d'axiomes insuffisamment formalisés, Hilbert proposa de rédiger une théorie purement formelle des mathématiques et de démontrer sa consistance *formelle*. En effet, on avait jusque-là toujours démontré les propriétés de consistance par rapport à des modèles : la géométrie de Bolyai-Lobachevsky avait été ramenée à la géométrie euclidienne, la géométrie euclidienne à l'analyse (Descartes)... Or ce type de démonstration montre seulement qu'une théorie est consistante si une autre théorie l'est. La propriété de consistance formelle demande simplement qu'à l'intérieur de la théorie formelle, on ne puisse avoir une formule qui soit démontrable et dont la négation soit démontrable (ce qui constituerait un paradoxe). Hilbert imposa également que les méthodes de démonstration de ce type de propriété soient "intuitivement convaincantes" et *finitistes*, c'est-à-dire ne faisant jamais appel

---

1. On confond parfois intuitionnisme et constructivisme, qui sont deux notions fort proches.  
2. Brouwer qui fut le fondateur de l'école intuitionniste.  
3. Rappelons que le principe du tiers-exclus dit que pour tout  $P$ ,  $P \vee \neg P$  est vraie.

à l'infini achevé. Hilbert appela cette nouvelle branche des mathématiques *métamathématique*. Hilbert énonça ce programme en 1908, et la réalisation commença vers 1920.

Hilbert se justifia de son attitude face à celle des intuitionnistes de la façon suivante : il scinda les mathématiques en deux domaines, celui des énoncés réels qui ont un sens intuitif et celui des énoncés idéaux qui n'en ont pas. Les énoncés idéaux ne sont que des moyens théoriques et élégants de simplifier les démonstrations d'une théorie donnée (cf. [Péter]). Il est ainsi plus facile de démontrer que  $a.n + b$  prend une infinité de valeurs premières si  $a$  et  $b$  sont premiers entre eux en utilisant des méthodes analytiques (théorie "idéale" des fonctions holomorphes) qu'en passant par une méthode "intuitive" n'utilisant que des résultats sur les nombres entiers. Il est même des cas où le recours à la théorie idéale est obligatoire, simplement parce qu'on ne connaît pas d'autres moyens de résoudre le problème. On peut également citer comme exemple concluant de cette attitude la construction de la géométrie projective (théorie idéale) à partir de la géométrie euclidienne traditionnelle.

Si l'ensemble de la communauté scientifique ne fut pas convaincu, nombre de mathématiciens se lancèrent cependant dans la réalisation du programme de Hilbert<sup>4</sup> et les résultats qu'ils allaient obtenir devaient se révéler assez inattendus.

## 5.4. L'arithmétique formelle

L'idée originelle de Hilbert était de construire un système formel des mathématiques, ou tout au moins de l'analyse. Mais le problème se révéla tellement complexe que lui et ses élèves se limitèrent à construire un système formel de l'arithmétique. Nous allons décrire ce système. Nous insistons cependant sur le fait que les définitions données ci-dessous n'ont pas toujours toute la précision nécessaire, mais la mise en place parfaitement rigoureuse d'un système formel de l'arithmétique dépasserait, et de beaucoup, le cadre de cet ouvrage.

L'arithmétique formelle est une extension de la théorie de la démonstration<sup>5</sup> du calcul des prédicats

**Définition 5.1 - Symboles** - *Les symboles de l'arithmétique formelle seront*

$$\leftrightarrow, \rightarrow, \wedge, \vee, \forall, \exists, =, +, \cdot, ', 0, a, b, c, \dots, |, (, )$$

Les lettres  $a, b, c, \dots$  sont les variables de notre langage. Afin d'en disposer d'une infinité, nous admettons également comme variable les termes de la forme  $a_{| \dots |}$ <sup>6</sup>. Nous allons définir les termes de notre langage.

**Définition 5.2 - Termes** -

1.  $0$  est un terme ;
2. les variables sont des termes ;
3. si  $r$  et  $s$  sont des termes alors  $(r)'$ ,  $(r) + (s)$ ,  $(r) \cdot (s)$  sont des termes.

Tout terme est obtenu en appliquant un nombre fini de fois les règles précédentes.

Il nous reste à définir les formules de notre langage :

**Définition 5.3 - Formules** -

1. si  $r$  et  $s$  sont des termes alors  $(r = s)$  est une formule ;
2. si  $A$  et  $B$  sont des formules alors  $(a \leftrightarrow B)$ ,  $(A \rightarrow B)$ ,  $(A \wedge B)$ ,  $(A \vee B)$  et  $(\neg A)$  sont des formules ;
3. si  $A$  est une formule et  $x$  une variable alors  $(\forall xA)$  et  $(\exists xA)$  sont des formules.

---

4. Notons que ce programme ne se résumait pas à la démonstration de la consistance de l'arithmétique formelle. Il comprenait 23 points, dont le premiers était précisément l'étude de l'hypothèse du continu de Cantor. Certains points du programme de Hilbert ne sont pas encore résolus aujourd'hui.

5. Une théorie formelle est une théorie de la démonstration et non une théorie des modèles.

6. Ainsi,  $a_{| | | |}$  représente formellement ce que l'on note généralement  $a_4$ .

Toute formule est obtenue en appliquant un nombre fini de fois les règles précédentes.

Munis du langage, nous allons pouvoir définir les axiomes et règles d'inférence du système formel  $\mathbb{N}$ .

**Définition 5.4 - Axiomes et règles d'inférence standards** - Les axiomes et les règles d'inférence du calcul des prédicats s'appliquent dans le système formel  $\mathbb{N}$ .

Les axiomes que nous allons définir ensuite sont appelés *axiomes non-logiques*. Pourquoi ce nom? Les axiomes que nous avons définis jusqu'ici ne remettent pas en cause les notions du calcul des prédicats et en particulier la notion de conséquence valide. En revanche, les axiomes que nous allons introduire ne peuvent plus s'interpréter au niveau de la notion de conséquence valide que comme des hypothèses de calcul, dans la mesure où ils n'ont pas de vérité logique. C'est en raison de ces axiomes que les résultats du calcul des prédicats ne peuvent s'étendre directement à l'arithmétique formelle. Nous essaierons de donner de chacun de ces axiomes une interprétation matérielle intuitive.

**Axiome 5.1 - Récurrence** -

$$A(0) \wedge \forall x(A(x) \rightarrow A(x')) \rightarrow A(x)$$

Si nous interprétons l'opérateur ' comme *successeur de*, alors l'axiome précédent se comprend immédiatement comme la définition de la récurrence : si la propriété est vraie pour 0 et si, quand elle est vraie pour  $n$ , elle est vraie pour  $n + 1$ , alors elle est vraie pour tout  $n$ <sup>7</sup>.

**Axiome 5.2 - Quatrième axiome de Peano** -

$$a' = b' \rightarrow a = b$$

Cet axiome formalise que :  $a + 1 = b + 1 \rightarrow a = b$

**Axiome 5.3 - Cinquième axiome de Peano** -

$$\neg a' = 0$$

Il n'existe aucun nombre ayant 0 pour successeur.

**Axiome 5.4 - Pseudo-transitivité de l'égalité** -

$$a = b \rightarrow (a = c \rightarrow b = c)$$

**Axiome 5.5 - Egalité**

$$a = b \rightarrow a' = b'$$

Axiome fondateur de l'égalité et de la notion de successeur avec le quatrième axiome de Peano ; si  $a = b$  alors  $a + 1 = b + 1$ .

**Axiome 5.6 - 0 élément neutre pour l'addition** -

$$a + 0 = a$$

**Axiome 5.7 - Successeur et addition** -

$$a + b' = (a + b)'$$

**Axiome 5.8 - 0 élément absorbant de la multiplication** -

$$a.0 = 0$$

---

7. Notons qu'ici les variables sont prises dans le sens de la généralité, comme dans toute suite du jeu d'axiomes. On pourrait donc tout aussi bien remplacer chaque axiome par sa clôture universelle.

Définit 0 comme élément absorbant de la multiplication.

### Axiome 5.9 - Distributivité de la multiplication -

$$a.b' = a.b + a$$

A partir des définitions données ci-dessus, il est possible de définir la suite des nombres 1,2,3... récursivement en posant  $1 = 0', 2 = 1' = 0'' \dots$

On ne peut définir dans  $\mathbb{N}$  que des fonctions polynomiales, en raison du symbolisme extrêmement pauvre du langage. Mais l'on peut, en revanche, définir pour toute fonction arithmétique  $f(x_1, \dots, x_n)$  un prédicat  $p(x_1, \dots, x_n, y)$  qui est vrai si et seulement si  $f(x_1, \dots, x_n) = y$ . On peut alors "parler" d'une fonction et opérer sur elle en travaillant sur le prédicat associé qui, lui, appartient bien au système formel. Ainsi considérons la fonction indicatrice<sup>8</sup> des nombres pairs  $P(x)$  et son prédicat associé  $p(x, y)$ . Le théorème  $P(x + 2) = P(x)$  s'exprimera dans  $\mathbb{N}$  par  $\exists u \exists v (p(x + 2, u) \wedge p(x, v) \wedge (u = v))$ .

Il est également possible de définir les symboles traditionnels de l'arithmétique classique comme synonymes de formules plus complexes. C'est le cas de  $>$  ou de  $<$  ou de l'exponentiation. En fait, l'essentiel des relations familières au mathématicien sont définissables.

On pense donc que le système formel  $\mathbb{N}$  que nous venons de définir a bien comme interprétation standard l'arithmétique telle que nous la pratiquons couramment. Si tel n'était pas le cas, il est clair que notre formalisme n'atteindrait pas son but.

Nous ne donnerons pas d'exemple de démonstration formelle dans  $\mathbb{N}$ . De telles démonstrations sont excessivement lourdes. La démonstration du théorème  $a = a$  ne prend pas moins de 17 lignes! Nous utiliserons d'ailleurs tout au long des paragraphes suivants des démonstrations informelles.

## 5.5. Décidabilité, complétude et consistance

Nous devons bien nous rappeler le but d'Hilbert : il s'agit, maintenant que les bases formelles sont posées, de démontrer la consistance formelle de  $\mathbb{N}$  (*i.e.* il n'existe pas dans  $\mathbb{N}$  une formule  $E$  telle que  $\vdash E$  et  $\vdash \neg E$ ).

En 1925, Ackermann publia une démonstration de consistance de  $\mathbb{N}$ . Mais deux ans plus tard, von Neumann montrait que la démonstration d'Ackermann ne s'appliquait qu'à un sous-système de  $\mathbb{N}$ , dans lequel l'axiome 5.1 est remplacé par un axiome voisin, mais plus faible.

En fait, les tentatives de démonstration de la consistance étaient vouées à l'échec comme allait le montrer Gödel. La démonstration originale de Gödel n'est pas celle que nous allons donner. Nous présenterons la démonstration donnée par Kleene en 1943, démonstration qui s'appuie sur les machines de Turing.

### 5.5.1. Décidabilité et calculabilité

Nous avons déjà effleuré les problèmes de calculabilité et de décidabilité dans le chapitre consacré aux machines de Turing.

Rappelons la définition que nous avons donnée de la calculabilité.

**Définition 5.5 - Calculabilité** - *Une fonction est dite calculable s'il existe une procédure algorithmique finie permettant de calculer en un nombre de pas fini sa valeur pour tout argument de son domaine.*

---

8. La fonction  $f$  indicatrice d'un ensemble  $E$  est la fonction définie par :  $\forall x \in E, f(x) = 1$  et  $\forall x \notin E, f(x) = 0$ .

La notion de décidabilité est très proche de la définition de la calculabilité, mais elle s'applique à des problèmes de décision *i.e.* à des classes de questions auxquelles la réponse est *oui* ou *non*.

*Une classe de questions est dite décidable s'il existe une procédure algorithmique finie permettant de résoudre n'importe laquelle des questions de cette classe en un nombre fini de pas.*

Des problèmes classiques de décision dans un système formel sont : "Une suite de symboles donnée est-elle une formule bien formée?" ou "Une suite d'expressions donnée est-elle la démonstration d'une formule?" ou encore "Une formule donnée est-elle démontrable?".

Les deux premières questions sont décidables. Il n'en est pas de même pour la troisième. Nous connaissons des cas où l'on sait résoudre le problème de décision : il s'agit par exemple du calcul propositionnel, par application de la méthode des tables de vérité. En revanche, nous ne savons pas a priori s'il existe une procédure de décision pour  $\mathbb{N}$ . Or une telle procédure permettrait de résoudre de façon simple et mécanique tous les problèmes concernant l'arithmétique élémentaire non encore résolus aujourd'hui. On pourrait ainsi prouver que la conjecture de Goldbach<sup>9</sup> ou le grand "théorème" de Fermat<sup>10</sup> sont démontrables ou indémontrables en appliquant mécaniquement cet algorithme. Il semble difficile de trouver une telle procédure. Il semblait encore plus difficile de prouver qu'elle n'existait pas. C'est pourtant ce que Church a fait.

La thèse de Church-Turing, que nous avons présentée au chapitre 10, ramène la notion de calculabilité intuitive à la notion de *calculabilité* pour les machines de Turing. Nous avons déjà évoqué les arguments en faveur de cette thèse. Nous la supposons admise dans la suite de l'énoncé. Rappelons la notion de *décidabilité* au sens de Turing : un prédicat est décidable au sens de Turing s'il existe une machine de Turing capable de déterminer sa vérité ou sa fausseté. La thèse de Church-Turing implique que la décidabilité intuitive et la décidabilité au sens de Turing sont équivalentes.

### 5.5.2. Construction d'une fonction non-calculable

Considérons de nouveau les machines de Turing que nous avons présentées au chapitre 4. Une machine de Turing est entièrement déterminée par son tableau, qui représente son programme. Ainsi le tableau 5.2 détermine totalement une machine de Turing.

	$B$	$0$	$1$
$z_0$	$(z_1, B, G)$	$(z_0, 0, D)$	$(z_0, 1, D)$
$z_1$	$(z_h, 1, I)$	$(z_h, 1, I)$	$(z_1, 0, G)$

Table 5.2 - Programme d'une machine de Turing

(celle qui calcule la fonction  $f(x) = x + 1$ ). Nous pouvons réécrire le programme précédent, en représentant un état  $z_i$  par  $i$ , sous la forme :

$$1, B, G, 0, 0, D, 0, 1, D; h, 1, I, h, 1, I, 1, 0, G.$$

Tout programme de machine de Turing peut donc s'écrire à l'aide des dix-sept symboles suivants<sup>11</sup> :

$$h \ 0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ B \ D \ G \ I \ , \ ;$$

Tout programme de machine de Turing peut donc être écrit sous la forme d'un nombre en base 17. De même, tout mot de  $\Gamma^*$  (de cardinal 11) peut se traduire de façon unique en un élément de  $\mathbb{N}$ .

9. Tout nombre pair supérieur ou égal à 4 est somme de deux nombres premiers.

10.  $x^n + y^n = z^n$  ne possède pas de solution entière pour  $n > 2$ . Andrew Wiles, de l'Université de Princeton, a proposé en juin 1993 une démonstration de cette conjecture.

11. Si on utilise un alphabet décimal. Nous pourrions nous contenter d'un système de numération binaire pour numéroter les états. Dans ce cas, neuf symboles suffiraient.

**Définition 5.7 - Index d'une machine de Turing** - Le nombre  $i$  (en base 17) qui représente une machine de Turing est appelé index de la machine de Turing, machine que nous noterons  $\mathcal{M}_i$ <sup>12</sup>.

Il découle immédiatement des résultats précédents le théorème :

**Théorème 5.1** - *L'ensemble des machines de Turing est dénombrable.*

Nous devons commencer à sentir le résultat plus proche, En effet, nous avons démontré au début de ce chapitre que l'ensemble des fonctions à variable entière était non-dénombrable, et nous venons de montrer que l'ensemble des fonctions que nous pouvons calculer est "seulement" dénombrable. Nous sommes donc déjà convaincus qu'il existe bien des fonctions non-calculables. Il nous suffit maintenant d'en exhiber une. Il nous reste quelques étapes à franchir avant d'en arriver là.

**Définition 5.8** - *Nous posons  $T(i, a, n)$  le prédicat qui prend la valeur vraie lorsque la machine de Turing dont l'index est  $i$ , appliquée à l'argument  $a$ , calcule au bout de  $n$  étapes exactement un résultat que nous noterons  $\varphi_i(a)$ .*

Remarquons que  $T(i, a, n)$  est également faux si  $i$  n'est pas l'index d'une machine de Turing valide. Nous allons établir deux résultats fondamentaux sur  $T$  et  $\varphi$ .

**Théorème 5.2** - *Le prédicat  $T(i, a, n)$  est décidable (au sens intuitif et au sens de Turing).*

Les valeurs de  $i$ ,  $a$  et  $n$  étant données, il est possible de vérifier que  $i$  est bien l'index d'une machine de Turing valide. Si ce n'est pas le cas,  $T$  est faux. Si c'est le cas, nous pouvons appliquer la machine de Turing  $\mathcal{M}_i$  à  $a$  et regarder si au bout de  $n$  étapes la machine vient juste de terminer le calcul d'une valeur. Si c'est le cas  $T(i, a, n)$  est vrai, sinon il est faux.

$T$  étant décidable au sens intuitif, il est décidable au sens de Turing, d'après la thèse de Church-Turing.

**Théorème 5.3** - *La fonction partielle  $\varphi_i$  est calculable.*

Nous disons que  $\varphi_i$  est une fonction *partielle* car elle n'est définie que pour les valeurs de  $a$  et de  $i$  telles qu'il existe un  $n$  tel que  $T(i, a, n)$  est vrai. Dire que  $\varphi_i$  est une fonction partielle calculable consiste simplement à dire que si  $\varphi_i$  est définie pour l'index  $i$  et l'argument  $a$ , *alors* elle est calculable. Le résultat dans ces conditions est évident. Il suffit d'appliquer la machine  $\mathcal{M}_i$  à l'argument  $a$  et attendre qu'elle ait calculé un résultat (ce qui arrivera car  $\exists n T(i, a, n)$  est vrai).

Il est temps pour nous d'arriver au résultat fondamental de cette section :

**Théorème 5.4** - *Soit la fonction  $\psi$  définie par :*

$$\psi(a) = \begin{cases} \varphi_a(a) + 1 & \text{s'il existe } n \text{ tel que } T(a, a, n) \\ 0 & \text{sinon} \end{cases}$$

*$\psi$  n'est pas calculable.*

Supposons en effet que  $\psi$  soit calculable. Alors il existe une machine de Turing  $\mathcal{M}_j$  qui calcule  $\psi$  et en particulier pour  $a = j$ , nous avons :

$$\psi(j) = \varphi_j(j)$$

Comme  $\mathcal{M}_j$  calcule  $\psi$ , nous avons pour tout  $a, T(j, a, n)$ . En particulier, pour  $a = j$ , nous savons qu'il existe  $n$  tel que  $T(j, j, n)$ . Or en appliquant directement la définition de  $\psi$  nous obtenons :

$$\psi(j) = \varphi_j(j) + 1$$

---

<sup>12</sup>. Notons qu'il existe des index qui ne désignent pas des machines de Turing valides. Cela est sans importance pour la suite.

Ce qui est en contradiction avec l'équation précédente. Q.E.D.

On remarquera que cette démonstration s'appuie une fois de plus sur la méthode de la diagonale cantorienne.

Il existe un corollaire important à ce théorème

**Théorème 5.5** - *Le prédicat  $\exists nT(a, a, n)$  est indécidable.*

Si ce prédicat était décidable, nous pourrions pour un  $a$  donné décider il existe  $n$  tel que  $T(a, a, n)$ . Si tel n'était pas le cas nous saurions que la valeur de  $\psi(a)$  est 0. Sinon, nous pourrions appliquer la machine de Turing à l'argument  $a$ , récupérer le résultat du calcul au bout des  $n$  étapes, lui ajouter 1, ce qui nous donnerait la valeur de  $\psi(a)$ .

Nous venons de décrire là une procédure<sup>13</sup> de calcul de  $\psi$ . Or  $\psi$  n'est pas calculable, donc par réduction à l'absurde,  $\exists nT(a, a, n)$  est indécidable. Plus généralement, le prédicat  $T(i, a, n)$  est indécidable. On en conclut que le problème de l'arrêt d'une machine de Turing est, lui aussi, indécidable.

### 5.5.3. Indécidabilité de $\mathbb{N}$

La démonstration de l'indécidabilité de  $\mathbb{N}$  va s'appuyer sur le résultat précédent. Il semble clair que le prédicat  $\exists nT(a, a, n)$  peut se formaliser dans  $\mathbb{N}$ . Les machines de Turing ne comprennent dans leur définition que des opérations arithmétiques élémentaires et le prédicat  $T$  ne fait appel à aucune notion autre qu'arithmétique ou liée au calcul des prédicats. Il doit donc exister dans  $\mathbb{N}$  une formule qui a comme interprétation<sup>14</sup>  $\exists nT(a, a, n)$ . Appelons  $C_a$  cette formule.

Remarquons maintenant que

$$\exists n T(a, a, n) \quad \text{entraîne} \quad \{ \vdash C_a \text{ dans } \mathbb{N} \}$$

En effet, si l'on sait qu'il existe un  $n$  tel que  $T(a, a, n)$ , on peut le démontrer informellement en construisant chacune des étapes de la machine de Turing associée. Une telle démonstration peut évidemment se formaliser dans  $\mathbb{N}$ .

Si nous supposons que seules les formules vraies sont démontrables<sup>15</sup>, nous avons

$$\{ \vdash C_a \text{ dans } \mathbb{N} \} \quad \text{entraîne} \quad \exists n T(a, a, n)$$

En fait,  $C_a$  représente exactement  $T(a, a, n)$ .

**Théorème 5.6 - Indécidabilité de  $\mathbb{N}$**  -  $\mathbb{N}$  est indécidable *i.e.* il n'existe pas de procédure de décision permettant d'établir la démontrabilité d'une formule.

---

13. Il faut tout de même ajouter une précision : la procédure décrite ci-dessus n'est correcte que lorsque l'on sait *construire* par une machine de Turing la machine de Turing calculant  $\varphi_a$  à partir de l'index  $a$  (ce qui est le cas au vu de notre définition de la fonction calculant l'index : celle-ci est clairement inversible). Dans le cas contraire, il faudrait inclure dans le programme de la machine calculant  $\psi$  tous les programmes de toutes les machines calculant tous les  $\varphi_a$ , ce qui est impossible, une machine de Turing devant avoir un nombre fini d'instructions.

14. Ici se trouve un des points principaux de la démonstration. C'est parce qu'on parle ici d'interprétation qu'on parlera ensuite de vérité de  $C_a$  si  $\exists nT(a, a, n)$  est vraie. Il doit y avoir adéquation de  $\mathbb{N}$  avec l'interprétation que l'on en fait : si  $\exists nT(a, a, n)$  est vraie, alors  $C_a$  est vraie sans qu'il soit encore question de savoir si elle est démontrable : elle est vraie parce que nous supposons que notre modèle est adéquat. Il faut faire bien attention de ne pas confondre *vrai* et *démontrable*.

15. Nous ne pouvons que le supposer, une telle propriété est, comme nous allons le voir, indémontrable par des méthodes métamathématiques. On peut cependant penser qu'il s'agit d'une condition indispensable pour la consistance de notre système formel. D'autre part, on peut remarquer que dans l'interprétation classique de l'arithmétique, les axiomes de  $\mathbb{N}$  sont vrais et l'application d'une règle d'inférence sur des axiomes ou des formules vraies donne systématiquement une formule vraie. Ce qui précède n'est cependant pas une démonstration acceptable car elle fait appel à des méthodes non finitistes, puisqu'elle considère l'ensemble achevé des démonstrations, qui est infini.

En effet, si une telle procédure existait, il serait possible, pour  $a$  donné, de savoir si  $C_a$  est démontrable ou non. On disposerait donc d'une procédure de décision pour  $T(a, a, n)$ , ce qui est impossible.

Il n'existe donc aucune procédure algorithmique permettant de déterminer la vérité ou la fausseté d'une formule quelconque de  $\mathbb{N}$  comme nous pouvions le faire avec les tables de vérité en calcul propositionnel.

#### 5.5.4. Incomplétude de $\mathbb{N}$

Nous allons avoir besoin de quelques propriétés supplémentaires pour démontrer l'incomplétude de  $\mathbb{N}$  (rapelons qu'il s'agit de montrer qu'il existe une formule  $F$  telle que l'on ait ni  $\vdash F$ , ni  $\vdash \neg F$ ).

Notons tout d'abord que comme dans  $\mathbb{N}$  seules les formules vraies sont démontrables, nous avons :

$$\{\vdash \neg C_a \text{ dans } \mathbb{N}\} \quad \text{entraîne} \quad \neg \exists n T(a, a, n)$$

Nous allons maintenant nous appuyer sur deux éléments fondamentaux liés à l'architecture de  $\mathbb{N}$ . Le premier est que dans  $\mathbb{N}$ , on peut reconnaître la démonstration d'une formule comme telle : la reconnaissance d'une démonstration est un problème décidable, comme nous l'avions déjà fait remarquer. Ensuite, nous devons remarquer que les symboles permettant de construire les formules de  $\mathbb{N}$  étant en nombre fini, le nombre de formules et de démonstrations est donc dénombrable, et l'on peut en établir une énumération. Ce dernier point est la clé de voûte de l'ensemble des démonstrations d'incomplétude dans les systèmes formels ; toute démonstration d'incomplétude revient finalement à montrer qu'il y a un nombre non dénombrable de formules vraies, alors qu'il y a seulement un nombre dénombrable de démonstrations (de la même façon que l'on montre qu'il y a une infinité non-dénombrable de fonctions et seulement un nombre dénombrable de fonctions calculables).

Les deux points précédents étant établis, il devient évident qu'il est possible de construire une machine de Turing  $\mathcal{M}_j$  qui, parcourant l'ensemble des démonstrations de  $\mathbb{N}$ , imprime 1 et s'arrête si elle trouve une démonstration de  $\neg C_a$ , et ne s'arrête jamais sinon. Ceci se résume par :

$$\exists n T(j, a, n) \quad \text{est équivalent à} \quad \{\vdash \neg C_a \text{ dans } \mathbb{N}\}$$

Nous pouvons donc résumer les trois points que nous avons établis et qui vont nous permettre de poursuivre la démonstration :

1.  $\{\vdash C_a \text{ dans } \mathbb{N}\}$  entraîne  $\exists n T(a, a, n)$  ;
2.  $\{\vdash \neg C_a \text{ dans } \mathbb{N}\}$  entraîne  $\neg \exists n T(a, a, n)$  ;
3.  $\neg \exists n T(j, a, n)$  est équivalent à  $\{\vdash \neg C_a \text{ dans } \mathbb{N}\}$ .

Nous pouvons maintenant démontrer le théorème d'incomplétude de Gödel. Supposons qu'il existe  $n$  tel que  $T(j, j, n)$ . Alors d'après (3), nous avons :  $\vdash \neg C_j$ . Mais alors (2) nous donne :  $\neg \exists n T(j, j, n)$ . Ceci contredit l'hypothèse. Donc nous avons  $\neg \exists n T(j, j, n)$ , donc  $\neg C_j$  est vrai.

Mais d'autre part,  $\neg \exists n T(j, j, n)$  impose d'après (3) que *non*  $\vdash \neg C_j$ .

Enfin, supposons que  $\vdash C_j$  ; alors d'après (1), nous savons qu'il existe un  $n$  tel que  $T(j, j, n)$ , ce qui impliquerait que  $C_j$  soit vrai. Or ceci est contradictoire avec le premier point que nous avons démontré. Donc, on n'a pas  $\vdash C_j$ .

Nous pouvons résumer nos trois résultats dans le théorème suivant :



**Théorème 5.7 - Théorème d'incomplétude** - Dans  $\mathbb{N}$ , nous pouvons trouver une formule liée  $C_j$  telle que

1. nous n'avons pas  $\vdash C_j$  ;
2. nous n'avons pas  $\vdash \neg C_j$  ;
1.  $\neg C_j$  est vraie.

Nous avons donc démontré que pourvu que  $\mathbb{N}$  soit adéquat et que seules les formules vraies soient démontrables dans  $\mathbb{N}$ ,  $\mathbb{N}$  est incomplet, c'est-à-dire qu'il existe des formules (vraies d'après l'interprétation standard) qui ne sont pas démontrables et dont la négation n'est pas non plus démontrable.

Essayons de comprendre comment nous sommes arrivés à ce résultat ; reprenons (3)

$$\exists n T(j, j, n) \text{ est équivalent à } \{\vdash \neg C_j \text{ dans } \mathbb{N}\}.$$

Dans l'interprétation que nous utilisons,  $\neg C_j$  exprime  $\neg \exists n T(j, j, n)$ . Nous avons donc construit une formule qui dans l'interprétation considérée exprime sa propre indémontrabilité. Il s'agit d'une assertion très proche du paradoxe du menteur.

Le résultat qui découle du théorème d'incomplétude est qu'il est impossible de formaliser l'ensemble des propriétés de l'arithmétique standard dans un système formel. La formule  $\neg C_j$  est une propriété de l'arithmétique, vraie, mais indémontrable. Une partie du programme de Hilbert s'écroulait.

### 5.5.5. Non démontrabilité de la consistance de $\mathbb{N}$

Nous ne donnerons qu'une idée de la démonstration montrant qu'il est impossible de démontrer la consistance formelle de  $\mathbb{N}$  dans le système formel  $\mathbb{N}$ .

Constatons tout d'abord que dans notre démonstration nous avons utilisé comme hypothèse :  $\{\vdash \neg C_a \text{ dans } \mathbb{N}\}$  entraîne  $\neg \exists n T(a, a, n)$ .

En fait cette hypothèse peut se réduire à la notion de consistance de  $\mathbb{N}$  : il n'existe pas dans  $\mathbb{N}$  de formule  $F$  telle que  $\vdash F$  et  $\vdash \neg F$ . Il suffit pour s'en convaincre de constater que  $(\{\vdash \neg C_a \text{ dans } \mathbb{N}\} \text{ entraîne } \neg \exists n T(a, a, n))$  se démontre à partir de  $(\neg \exists n T(a, a, n) \text{ entraîne } \{\vdash \neg C_a \text{ dans } \mathbb{N}\})$  et de la consistance simple.

Donc nous avons : ( $\mathbb{N}$  est simplement consistant) entraîne  $(\neg C_j \text{ est vrai})$  (ici  $j$  représente le  $j$  particulier du théorème précédent).

La suite de la démonstration consiste à remarquer que l'on peut exprimer de façon formelle la proposition :  $\mathbb{N}$  est simplement consistant. Il suffit de considérer que l'on peut construire une machine de Turing  $\mathcal{M}_k$  cherchant dans les démonstrations de  $\mathbb{N}$  une démonstration d'une formule de la forme  $F \wedge \neg F$ , qui imprime 1 si elle en trouve une et ne termine jamais sinon.

Alors la consistance de  $\mathbb{N}$  est équivalente au fait que, pour un  $a$  quelconque, il n'existe pas de  $n$  tel que  $T(k, a, n)$ . En particulier, il ne doit pas exister de  $n$  tel que  $T(k, k, n)$ . Ce prédicat peut être formalisé, nous le savons déjà, par une formule de  $\mathbb{N}$  que nous appellerons  $C$ . Nous pouvons alors reprendre la proposition : ( $\mathbb{N}$  est simplement consistant) entraîne  $(\neg C_j \text{ est vrai})$ . Elle devient formellement :

$$C \rightarrow \neg C_j$$

Hilbert et Bernays ont montré en 1939 que cette formule est démontrable dans  $\mathbb{N}$ , ce qui est assez naturel dans la mesure où  $\mathbb{N}$  doit être adéquat à notre modèle intuitif de l'arithmétique. Donc :

$$\vdash C \rightarrow \neg C_j.$$

Supposons maintenant que  $\vdash C$  (c'est-à-dire que la consistance de  $\mathbb{N}$  est démontrable), alors nous obtenons  $\vdash \neg C_j$ , ce qui contredit le théorème précédent (nous avons montré par le théorème d'incomplétude que nous n'avons pas :  $\vdash \neg C_j$ ).

D'où le résultat :

**Théorème 5.8 - Indémontrabilité de la consistance de  $\mathbb{N}$**  - La formule  $C$  exprimant formellement dans  $\mathbb{N}$  la consistance de  $\mathbb{N}$  est indémontrable dans  $\mathbb{N}$ .

## 5.6. Les conséquences

Nous ne nous attarderons pas sur les conséquences du théorème de non-démontrabilité de la consistance. Disons simplement que, pour nombre de mathématiciens, il signifia la fin de l'espoir d'obtenir une garantie de la sûreté des mathématiques classiques.

En revanche, la notion de non-décidabilité et de non-calculabilité nous intéresse au premier chef. Ces résultats ont donné naissance à une branche fructueuse de l'informatique, la théorie de la calculabilité et de la complexité. Ils ont également fourni un argument aux défenseurs de l'IA : "dans la mesure où des fonctions ne sont pas calculables, et où nous voulons tout de même obtenir des résultats utilisables, nous sommes obligés d'utiliser des méthodes heuristiques, méthodes caractéristiques d'une approche IA". Il faut cependant prendre ce raisonnement avec prudence ; un algorithme d'intelligence artificielle s'exécutera, comme tout autre algorithme, sur une machine de Turing, et est donc soumis aux mêmes limitations.

Le théorème d'incomplétude nous intéresse également dans la mesure où, selon certains, il montre qu'il existe des propositions vraies qu'aucun moyen mécanique ne permet de démontrer ; certains adversaires de l'IA se sont servis de cet argument pour prétendre qu'un calculateur ne pourrait jamais réaliser certains travaux réalisés par l'esprit humain. L'argument doit aussi être prudemment considéré.

On peut, d'autre part, se demander si la démonstration d'incomplétude ne vient pas du jeu d'axiomes que nous avons utilisé au départ. Si nous ne parvenons pas à démontrer une propriété que nous savons vraie (comme celle que nous avons construite lors de la démonstration du théorème d'incomplétude), il nous suffit de l'inclure comme axiome supplémentaire. On pourrait ainsi espérer construire un système complet par ajout successif de toutes les propriétés vraies que nous ne pouvons démontrer. Il existe malheureusement une objection majeure à cette méthode. Une extension du théorème de Gödel stipule que pour tout système formel contenant l'arithmétique, il est possible de construire une propriété vraie non-démontrable, en utilisant d'ailleurs une méthode très voisine de celle que nous venons de présenter. Ceci ruine définitivement nos espoirs. Disons tout de même que le théorème d'incomplétude n'est pas aussi anti-naturel ou anti-intuitif qu'il peut apparaître au premier abord. En fait, il exprime que notre modèle intuitif de l'arithmétique est trop riche pour être réduit à un quelconque système formel, ce qui peut sembler relativement normal. Depuis la démonstration du théorème d'incomplétude, des modèles différents vérifiant l'axiomatique formelle de l'arithmétique ont été développés ; il a été démontré que l'ensemble de toutes les formules vraies dans tous les modèles est exactement l'ensemble des théorèmes.

Au sujet des modèles non-standard, Thoralf Skolem a démontré, en étendant un théorème dû à Lowenheim, que toute formalisation axiomatique du premier ordre de la théorie des ensembles (dont le modèle standard admet, bien entendu, des ensembles non-dénombrables) admet un modèle dont le domaine d'objets est au plus dénombrable. Donc la théorie naïve des ensembles ne peut être formalisée de façon catégorique<sup>16</sup> par un système axiomatique du premier ordre. Ce théorème n'est pas aussi paradoxal qu'il peut le paraître. En effet, toute théorie axiomatique formelle contient au plus, nous l'avons vu, un nombre dénombrable de démonstrations. Il semble clairement difficile dans ces conditions de représenter des ensembles

16. Une théorie est catégorique ssi l'ensemble de ses modèles est isomorphe *i.e.* pour tout couple de modèles  $M_1, M_2$ , il existe une bijection  $f$  entre les domaines d'interprétation de ces deux modèles telle que pour tout symbole de prédicat  $P$  d'arité  $n$ ,  $M_1(P)(a_1, \dots, a_n)$  est vrai ssi  $M_2(P)(f(a_1), \dots, f(a_n))$  est vrai et telle que pour tout symbole fonctionnel  $g$  d'arité  $n$ ,  $M_1(g)(a_1, \dots, a_n)$  vaut  $a$  ssi  $M_2(g)(f(a_1), \dots, f(a_n))$  vaut  $f(a)$ . Tous les modèles d'une théorie catégorique sont donc mathématiquement indiscernables.

non-dénombrables. On peut aussi mettre le théorème de Lowenheim-Skolem en relation avec le théorème d'incomplétude : considérons un système  $S$  "riche". Comme tous les systèmes "riches" sont incomplets, on peut trouver une propriété  $p$  telle que les systèmes  $S \cup p$  et  $S \cup \neg p$  soient des systèmes "corrects". Ces deux systèmes ne peuvent être catégoriques car les interprétations ne sont pas isomorphes.

Remarquons enfin la grande différence qui existe entre le théorème d'incomplétude et le théorème énonçant l'impossibilité de démontrer la consistance de  $\mathbb{N}$  dans  $\mathbb{N}$ . Le premier stipule qu'il existe des propriétés vraies non démontrables. *Le second ne prétend en aucun cas qu'il existe des propriétés démontrables qui ne soient pas vraies.* Il dit simplement que l'on ne peut démontrer la consistance de  $\mathbb{N}$  en utilisant seulement les axiomes de  $\mathbb{N}$  et des méthodes finitistes, comme l'exigeait Hilbert. Une démonstration de consistance utilisant des méthodes non-finitistes a d'ailleurs été publiée par Gerhard Gentzen peu après que Gödel ait énoncé son théorème. Cette méthode satisfait de nombreux mathématiciens, même si elle fait appel à la notion d'induction transfinie. Il ne faut donc en aucun cas parler de théorème d'inconsistance comme on parle du théorème d'incomplétude mais bien d'un théorème de non-démonstrabilité finitiste de la consistance, ce qui est bien différent.

## 5.7. Faut-il jeter la logique ?

Les résultats obtenus par Gödel ont eu des conséquences très importantes sur la communauté des mathématiciens eux-mêmes. Certains d'entre eux, dont Imre Lakatos, ont même estimé nécessaire de renoncer à la théorie de la démonstration pour ne garder aux démonstrations mathématiques que leur caractère intuitif ([Lakatos])

*"Pourquoi ne disons-nous pas que le test ultime permettant de savoir si une méthode peut-être admise en arithmétique revient à la question de savoir si elle est intuitivement convaincante ?"*

Il faut cependant faire une distinction entre ce qui apparaît comme deux parties fondamentalement différentes de la logique, que l'on a parfois tendance à confondre :

- Il y a d'une part un problème à caractère philosophique qui s'interroge sur la valeur sémantique des axiomes logiques ; Est-ce qu'un axiome est vrai ? Qu'est-ce que la vérité ? Ce problème interpelle les philosophes anglo-saxons depuis près d'un siècle ; un bon livre faisant le point sur la philosophie de la logique est [Engel].
- Il y a d'autre part un problème mathématique qui s'interroge sur la relation entre une théorie de la vérité, et une axiomatisation de cette théorie. C'est essentiellement à ce point que nous nous sommes intéressés.

En fait, ces deux points correspondent à deux étapes pragmatiques du travail du mathématicien ou du logicien. Face à une réalité trop vaste pour être traitée efficacement, il commence par la réduire en un langage formel dans lequel il définit une théorie de la vérité et un modèle, puis il construit une théorie de la démonstration qui lui permet de vérifier la validité de ses intuitions dans le cadre de ce modèle. Les deux étapes sont indispensables, mais il ne faut pas attendre plus de la logique que ce qu'elle peut donner : sa démarche est réductionniste, et on ne peut trouver au bout de la chaîne un système aussi riche que le monde réel qu'elle modélise. La crise de la science actuelle est plus une crise liée aux limites de cette science qu'à ses fondements : ce qui est en cause, c'est le dogme *réductionniste* qui, si l'on caricature un peu, prétendait transformer *sans perte de généralité* le monde en un système axiomatique fini où tout serait démontrable. Il semble clair que le dogme a fait long feu, dans les deux domaines ; les philosophes semblent accepter l'impossibilité de réduire le langage ou la pensée humaine, ainsi que les notions intuitives de vérité et de validité sous-jacentes, à un système unique ; les mathématiciens sont conscients que la théorie sémantique et la théorie de la démonstration ne sont pas équivalentes, et qu'aucun système axiomatique fini ne peut représenter complètement et catégoriquement un modèle un peu complexe.

Il ne faut pourtant pas rejeter le travail accompli, ni refuser de le continuer. Il faut accepter la vision pragmatiste qui consiste à regarder la logique comme un outil, imparfait certes, mais pourtant indispensable. On ne peut rejeter les axiomes de la logique modale sous prétexte qu'elle est "née dans le péché" (Quine), si celle-ci fournit un modèle utile pour formaliser le raisonnement. De même, on ne peut rejeter la théorie de la démonstration, sous prétexte qu'elle ne fournit pas une certitude absolue : elle n'est qu'une vérification de nos intuitions et a rempli et continuera à remplir parfaitement ce rôle.

## Références bibliographiques

- [Engel] Pascal Engel. *La norme du vrai. Philosophie de la logique*. NRF Essais. Gallimard, 1989.
- [Lakatos] Imre Lakatos. *Preuves et réfutations*. Hermann, 1984.
- [Péter] Rózsa Péter. *Jeux avec l'infini : voyage à travers les mathématiques*. Editions du seuil, collection Points Sciences, 1977.  
L'édition hongroise originale date de 1957. ISBN : 2-02- 004568-0.