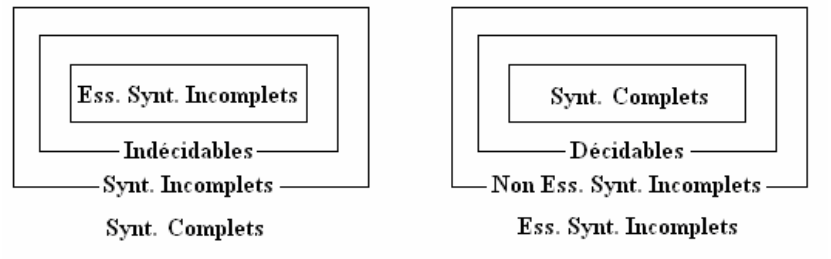
La Tétralogique.

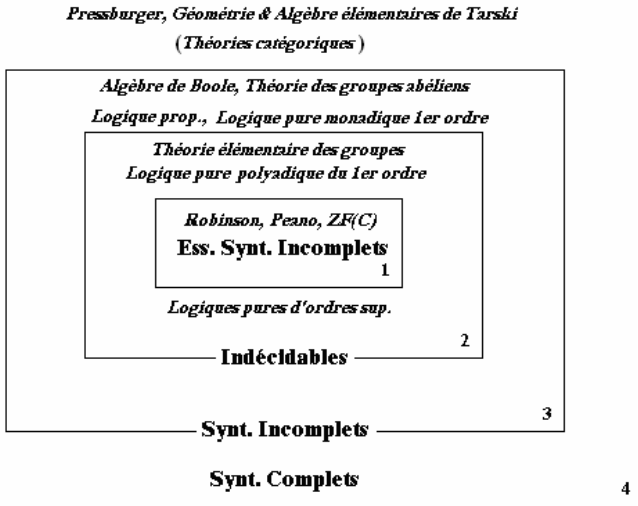
III. L'universalité au sens de Gödel.

Incomplétude syntaxique et indécidabilité algorithmique.

Nous avons défini, dans la première partie de ce travail de synthèse, quelques caractéristiques importantes des systèmes formels en nous concentrant particulièrement sur les notions d'incomplétude syntaxique et d'indécidabilité algorithmique. Nous avons situé ces propriétés dans les diagrammes, C-D, complémentaires suivants :



Le moment est venu de placer à leur place exacte quelques systèmes couramment utilisés. Nous montrons dans cette troisième partie que le diagramme de gauche doit être meublé comme suit. Le lecteur transposera sans peine s'il préfère le diagramme de droite.



Les systèmes situés en zone 4 incarnent l'idéal Hilbertien : ils sont syntaxiquement complets donc décidables. Les théories d'un seul modèle, dites catégoriques, en font naturellement partie, en particulier, l'algèbre et la géométrie élémentaires. La zone 3 regroupe les systèmes syntaxiquement incomplets mais décidables. La zone 2 comprend les systèmes indécidables et syntaxiquement incomplets quoique non essentiellement syntaxiquement incomplets.

Les systèmes les plus puissants sont situés en zone 1 : ils sont essentiellement syntaxiquement incomplets donc indécidables. La théorie des nombres se trouve en zone 1 à égalité avec la toute puissante théorie des ensembles! On pense que le seuil d'universalité au sens de Gödel, à partir duquel tout système est capable d'émuler n'importe quel autre système, y compris lui-même, est proche de la frontière qui limite la zone 1.

Nous proposons de passer en revue les quatre zones dans l'ordre suivant, 3, 2, 4 et enfin, 1. Nous commençons par les zones, 3 et 2, qui contiennent les fragments et extensions usuels de la logique binaire classique. La logique est la branche des mathématiques qui a pour ambition de codifier les raisonnements valides et il semble naturel de commencer par elle. On pourrait penser qu'une telle entreprise est de tout repos dans la mesure où il semblerait qu'une bonne dose de bons sens devrait permettre de régler cette question. Il n'en n'est rien et les traités de logique sont épais et complexes sans qu'il y ait apparemment moyen de faire autrement. Tout au plus pourrait-on espérer que les logiciens fassent davantage d'efforts dans l'uniformisation de leur vocabulaire et dans la clarté de leur présentation.

Nous poursuivrons par l'étude de la zone, 4, qui comprend les variantes élémentaires des mathématiques usuelles, algèbres et géométries, telles qu'elles sont plus ou moins enseignées à l'école secondaire. Nous verrons cependant que des variantes non élémentaires existent et que leur position dans le diagramme, C-D, dépend largement du cadre dans le quel on développe ces disciplines.

Enfin, nous terminerons par les systèmes les plus puissants qui occupent la zone, 1. Nous dirons quelques mots de la toute puissante théorie des ensembles puis nous montrerons que, contrairement à toute attente, l'arithmétique révèle un pouvoir d'expression aussi puissant.

Un système situé en zone 3 : La logique des propositions.

La logique propositionnelle est le fragment minimal de la logique classique. Fondée par le logicien allemand Frege, elle définit les lois formelles du raisonnement valide sur les énoncés, dits clos, qui ne dépendent d'aucun paramètre et pour les quels il fait sens de parler de vérité ou de fausseté. "7 est un entier premier" ou "3 est un entier pair" sont deux propositions acceptables en logique propositionnelle. Cette logique ne permet pas, loin de là, d'exprimer toutes les propositions mathématiques que l'on a généralement en vue. Par exemple, la proposition, "n est un entier premier", n'y est pas exprimable car elle contient une variable libre, n, dont on ne sait rien. Malgré son caractère rudimentaire, la logique propositionnelle constitue un point de départ commode pour apprendre à maîtriser les concepts de base relatifs aux systèmes formels. Elle est construite sur base des éléments suivants.

Alphabet de base :
p, q, r, …, un ensemble de, n, propositions simples, encore dites atomiques, pour lequel il fait sens de parler de vérité, deux symboles, V et F respectivement pour la tautologie et la contradiction, un symbole, $\neg$, pour la négation, quelques connecteurs bien choisis parmi une liste détaillée ci-après qui en comprend dix et enfin un jeu de parenthèses pour éliminer les ambiguïtés dans les notations.

Grammaire (régulière) de formation :
$S \to \{V, F, p, q, r, \ldots\}$; $S \to \neg S$; $S \to SOS$, $O \to \{\wedge, \vee, \Rightarrow, \Leftrightarrow\}$.
(Un jeu d'accolades signifie qu'un seul symbole doit y être prélevé à la fois).

Voici une proposition bien formée : $((\neg p \vee q) \wedge r) \vee ((p \Rightarrow q) \Rightarrow r)$.

Axiomatisation standard :
On axiomatise habituellement la logique propositionnelle sur base des connecteurs logiques $\{\neg, \wedge, \vee\}$, plus l'égalité, soit dans la notation de Wolfram :

$$p \wedge q = q \wedge p \quad p \vee q = q \vee p$$
$$p \wedge (q \vee \neg q) = p \quad p \vee (q \wedge \neg q) = p$$
$$p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r) \quad p \vee (q \wedge r) = (p \vee q) \wedge (p \vee r)$$

Il est possible d'axiomatiser le même système de façon plus courte en n'utilisant que le seul connecteur logique, $Nand = \overline{\wedge}$. On a alors :

Axiomatisation courte :

$$p \overline{\wedge} (p \overline{\wedge} q) = p \overline{\wedge} (q \overline{\wedge} q)$$
$$p \overline{\wedge} (p \overline{\wedge} (q \overline{\wedge} r)) = q \overline{\wedge} (q \overline{\wedge} (p \overline{\wedge} r))$$

Il existe même, sur base de ce même connecteur, une axiomatisation minimale et on peut montrer qu'on ne peut pas faire plus concis :

Axiomatisation minimale :

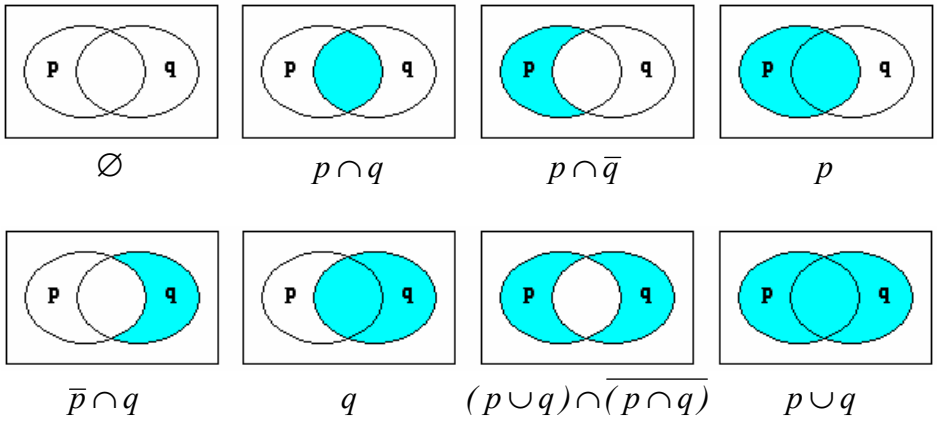
$$((p \overline{\wedge} q) \overline{\wedge} r) \overline{\wedge} (p \overline{\wedge} ((p \overline{\wedge} r) \overline{\wedge} p)) = r$$

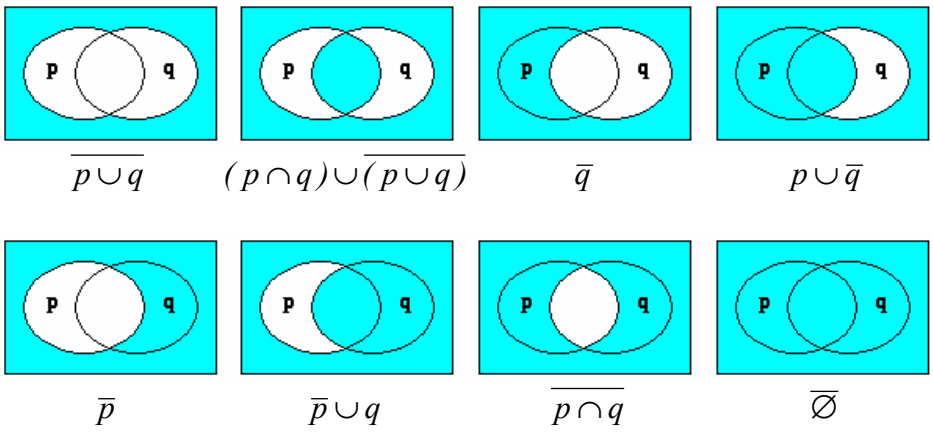
On voit que l'on perd progressivement en lisibilité ce que l'on gagne en concision.

L'axiomatisation standard est isomorphe ($\oplus \to \vee, \otimes \to \wedge, = \to \Leftrightarrow, \bar{} \to \neg$) à celle d'une algèbre finie de Boole, $(a,b,c) \in \{V,F\}$:

$$a \oplus b = b \oplus a \qquad a \otimes b = b \otimes a$$
$$a \oplus (b \otimes c) = (a \oplus b) \otimes (a \oplus c) \qquad a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$$
$$a \oplus \bar{a} = V \qquad a \otimes \bar{a} = F \qquad a \oplus F = a \qquad a \otimes V = a$$

De même le minuscule fragment de la théorie des ensembles finis réduit à l'union, l'intersection et la complémentation accepte la même axiomatisation ($\oplus \to \cup, \otimes \to \cap, \bar{} \to Compl, F \to \varnothing, V \to \overline{\varnothing}$):





Interprétations et modèles.
La logique propositionnelle portant sur n propositions, {p, q, r, …}, possède exactement 2^n modèles non isomorphes qui se différencient par leur table de vérité. Chaque variable est interprétée comme une proposition close et chaque connecteur reçoit son interprétation habituelle : $\wedge$ = "et", $\vee$ = "ou", $\Rightarrow$ = "implique", et enfin, $\Leftrightarrow$ = "est équivalent à". Le tableau suivant épuise toutes les possibilités dans le cas, n=2, et il est immédiatement généralisable lorsque le nombre de variables croît.

F	$p \wedge q$	$p > q$	p	$p < q$	q	$p \oplus q$	$p \vee q$	$p \overline{\vee} q$	$p \Leftrightarrow q$	$\neg q$	$p \Leftarrow q$	$\neg p$	$p \Rightarrow q$	$p \overline{\wedge} q$	V
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

Dans ce cas particulier, n=2, il existe donc quatre univers d'interprétation possibles, celui où p et q sont vraies, celui où p et q sont fausses et ceux où l'une est vraie et l'autre fausse. Rappelons qu'une proposition donnée est valide si elle est vraie dans toutes les interprétations. Par exemple, la proposition suivante est valide :

$$p \overline{\wedge} q = q \overline{\wedge} p .$$

La logique propositionnelle est correcte, cohérente et sémantiquement complète car toute proposition valide est prouvable sur base de ses axiomes. On peut comprendre intuitivement pourquoi il en est ainsi. On démontre que toute formule propositionnelle, P, peut toujours être ramenée, au choix, à l'une ou l'autre des formes normales équivalentes, conjonctives ou disjonctives, soit respectivement,

$$P \Leftrightarrow \bigwedge_{i=1}^{m} \left(\bigvee_{j=1}^{n} Q_{i,j} \right) \qquad \text{ou} \qquad P \Leftrightarrow \bigvee_{i=1}^{m} \left(\bigwedge_{j=1}^{n} Q_{i,j} \right)$$

Par exemple, on a que les formes suivantes sont équivalentes :

$$p \wedge (\neg q \vee r) \wedge (p \vee \neg r) \Leftrightarrow (p \wedge r) \vee (p \wedge \neg q) .$$

En Mathematica, c'est la fonction, LogicalExpand[P], qui calcule la forme normale disjonctive de P. Aucune fonction n'est prévue pour le calcul direct de la forme conjonctive et il y a lieu de passer par les lois de De Morgan,

$$\neg\left(\bigvee_{j=1}^{n} Q_{i,j}\right) \Leftrightarrow \left(\bigwedge_{j=1}^{n} \neg Q_{i,j}\right) \qquad \text{et} \qquad \neg\left(\bigwedge_{j=1}^{n} Q_{i,j}\right) \Leftrightarrow \left(\bigvee_{j=1}^{n} \neg Q_{i,j}\right).$$

C'est d'autant plus étrange que la forme conjonctive, qui représente une sorte de factorisation booléenne, est la plus utile. C'est en tous cas celle que l'on utilise dans l'argument de complétude sémantique suivant. Etant donnée une formule valide, P, chacun des termes, $Q_{i,j}$, qui composent sa forme normale conjonctive doit impérativement contenir une paire, $q \vee \neg q$. Or, $q \vee \neg q$ est facilement prouvable à partir des axiomes de base de la logique propositionnelle, d'où il résulte que chaque terme, $Q_{i,j}$, est prouvable et qu'enfin c'est la proposition, P, tout entière qui l'est comme conjonction de termes prouvables.

La logique propositionnelle est syntaxiquement incomplète puisqu'elle possède plusieurs modèles non isomorphes. Par exemple, la proposition, $p = q$, ne peut pas être démontrée ni réfutée sur base des axiomes, c'est un indécidable du système. Cela est bien normal puisqu'il existe des univers où cette proposition est vraie et d'autres où elle est fausse.

La logique propositionnelle est syntaxiquement incomplète et cependant elle est décidable ce qui la situe en zone 3 dans le diagramme, C-D. Vu la complétude sémantique de la logique propositionnelle, une procédure de décision qui teste la validité donc la théorémicité d'une proposition donnée, P, consiste à vérifier qu'elle est vraie dans toutes ses interprétations. Comme le nombre des interprétations est fini, la procédure qui consiste à les essayer toutes est assurée de s'arrêter. Cette procédure est cependant terriblement inefficace puisqu'elle nécessite 2^n passages dans le plus mauvais cas : cent variables booléennes, ce qui n'aurait rien d'extravagant, pourraient requérir 10^{30} vérifications !

Les problèmes où n dépasse la centaine sont monnaie courante, des horaires de chemin de fer au sudoku. L'inefficacité de la méthode exhaustive est en relation avec la difficulté d'un problème plus général, le problème, SAT. Le problème SAT demande une procédure pour trouver un jeu de variables booléennes qui satisfont une expression logique imposée. Aucun algorithme n'est connu qui règle toutes les instances de ce problème en un temps polynomial. L'existence d'un tel algorithme signifierait que SAT appartiendrait à la classe P. Par contre, il est connu que ce problème appartient à la classe NP qui se définit comme l'ensemble des problèmes dont la solution est vérifiable en temps polynomial. Attention, sauf si, NP≠P, NP ne signifie pas "non polynomial" mais "non déterministe polynomial" ce qui signifie qu'il existe un algorithme non déterministe qui fonctionne en temps polynomial, autrement dit qui se contente de tirer au hasard une solution possible puis de la vérifier. Cet algorithme ne fonctionne évidemment que si l'on est très chanceux. La conjecture, NP≠P, est l'un des défis majeurs du XXI^{ième} siècle en informatique théorique. Le statut de sa prouvabilité, de sa réfutabilité ou de son indécidabilité sera envisagé à la fin de cet exposé. Sous réserve de l'hypothèse, NP≠P, il est exclu de trouver une procédure effective qui règle toutes les instances du problème SAT en un temps polynomial. Par contre, rien n'empêche qu'il existe des algorithmes qui en résolvent rapidement un grand nombre d'instances. Les performances des solutionneurs qui participent aux compétitions, organisées en parallèle avec les congrès consacrés au problème SAT, en attestent.

Mathematica a développé une stratégie de résolution du problème SAT au travers de la commande, LogicalExpand[]. Lorsque LogicalExpand[P], répond True, c'est que P est valide donc prouvable et lorsqu'elle répond False, c'est que P est insatisfaisable donc réfutable. Dans tous les autres cas, on ne peut rien affirmer : soit P est contingente donc indécidable, soit P est prouvable ou réfutable mais la procédure ne l'a pas détecté, suite à une stratégie défaillante.

Le cas de la logique propositionnelle est suffisamment simple pour qu'on ait été capable de construire des algorithmes de constructions de preuves. La longueur de ces preuves dépend évidemment de l'axiomatisation choisie, standard, courte ou minimale. Sans surprise, les preuves sont d'autant plus longues que le cadre axiomatique choisi est concis. Par exemple, pour prouver la proposition, $p \barwedge q = q \barwedge p$, dans l'axiomatisation courte, on n'a jamais réussi à faire plus court que 16 étapes,

$$p \barwedge q = p \barwedge ((q \barwedge q) \barwedge (q \barwedge q)) = p \barwedge (p \barwedge (q \barwedge q)) = \cdots = q \barwedge p$$

alors que dans la forme équivalente, $\neg(p \wedge q) = \neg(q \wedge p)$, elle découle quasi immédiatement des axiomes standards.

Tous les problèmes relatifs à la logique propositionnelle ne sont pas simples. S'il est facile de déduire les propositions suivantes de l'axiomatique standard :

$$p \vee q = q \vee p \qquad p \vee (q \vee r) = (p \vee q) \vee r \qquad \neg(\neg(p \vee q) \vee \neg(p \vee \neg q)) = p,$$

la réciproque, qui semble anodine, n'a été établie qu'en 1996 et en encore avec le secours d'un ordinateur. Elles constituent l'axiomatique de Robbins de la logique propositionnelle.

La logique propositionnelle n'a qu'un pouvoir d'expression limité. Elle est très loin de pouvoir exprimer toutes les propositions que l'on envisage habituellement en mathématiques. Même les simples syllogismes aristotéliciens ne rentrent pas dans son cadre. Cela tient au fait que si la proposition, p = "Jean est mortel", y est formulable, il n'en va pas de même de sa généralisation, p = "Tous les hommes sont mortels". Il manque en particulier la notion de variable qui autorise les substitutions du type, "homme→Jean". Le calcul des propositions n'est en fait qu'une première étape dans la formalisation du raisonnement logique.

Un système situé en zone 2 : La logique du premier ordre.

La logique du premier ordre élargit le cadre de la logique propositionnelle. A cette fin, elle définit :

- des constantes, V, F, a, b, c, …, en nombre quelconque qui désignent des objets individuels;
- des variables, x, y, z , …, en nombre quelconque qui portent sur une famille d'objets sans en désigner un particulièrement;
- des fonctions d'arité, n, portant sur n variables;
- des prédicats d'arité, n, qui sont autant de fonctions booléennes dont les n arguments peuvent être des constantes, des variables ou des fonctions;
- deux quantificateurs dits, existentiel, $\exists$, et universel , $\forall$, qui s'appliquent aux constantes ou aux variables mais pas aux fonctions ni au prédicats. Chacun lie la variable à laquelle il s'applique qui, de ce fait, devient muette dans une formule quantifiée;
- enfin les connecteurs logiques, $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$, etc … empruntés à la logique booléenne.

Remarque purement anecdotique : de même que les connecteurs booléens, $\neg$ et $\wedge$, se suffisent à eux-mêmes, le quantificateur, $\forall$, suffit également pour exprimer toutes les quantifications. On a, en effet, $\exists x : P(x) \Leftrightarrow \neg(\forall x : \neg P(x))$. Il n'est cependant pas dans les usages de se priver de $\forall$.

Axiomatique.

La logique du premier ordre étant une extension de la logique propositionnelle, elle en conserve tous les axiomes et en ajoute de nouveaux qui règlementent l'usage des variables et des quantificateurs. Ces axiomes additionnels s'écrivent, dans la notation de Wolfram,

$$x\_ \wedge y\_ \Rightarrow x\_ \qquad x\_ \Rightarrow \forall_{y\_} x\_ \qquad x\_ \Rightarrow x\_ \wedge \#\&$$

$$\forall_{a\_}(b\_ \Rightarrow c\_) \Rightarrow (\forall_{a\_} b\_ \Rightarrow \forall_{a\_} c\_)$$

$$a\_ \Rightarrow \forall_{b\_} a\_ \quad /; FreeQ[a,b]$$

$$\exists_{a\_} a\_ == b\_ \quad /; FreeQ[a,b]$$

$$a\_ == b\_ \Rightarrow (c\_ == d\_) \quad /; (FreeQ[c, \forall\_\_] \wedge MatchQ[d, c/.a\_ :> a \vee b])$$

FreeQ[expr,forme] teste si forme est présent quelque part dans expr, par exemple, *FreeQ[a,a+b]=False mais FreeQ[a+b,a]=True*. *MatchQ[expr,forme]* teste si expr respecte la structure formelle notée forme, par exemple, *MatchQ[a⇒b,_ ⇒_]=True mais MatchQ[a∧b,_ ⇒_]=False*.

Comme la logique propositionnelle, la logique du premier ordre est consistante, correcte et sémantiquement complète. La complétude sémantique est particulièrement importante car elle garantit que prouvabilité et validité y coïncident d'où l'existence assurée d'une bonne notion de preuve. Cette logique permettant d'exprimer l'essentiel de ce dont on a besoin en mathématiques, elle a été adoptée comme langage naturel par l'immense majorité des mathématiciens.

La logique du premier ordre est syntaxiquement incomplète comme l'était la logique propositionnelle mais, par contre, elle n'est plus, en toute généralité, que semi décidable ce qui la situe en zone 2 dans le diagramme, C-D.

On peut comprendre intuitivement pourquoi, en logique du premier ordre, il n'existe plus de procédure de décision pour la validité donc pour la théorémicité d'une proposition, P. Il n'est certainement plus possible de tester la validité d'une proposition donnée en tentant d'épuiser les tables de vérité correspondant à toutes les interprétations possibles car la présence de variables et de quantificateurs rend leur nombre infini. Quant à généraliser la fonction, LogicalExpand[], cela a été tenté, mais cette procédure étendue ne s'arrête en toute certitude que si P est valide sinon elle peut se mettre à boucler.

Voici deux exemples de propositions que l'on peut soumettre à LogicalExpand[] qui répond que la première est trivialement valide et que l'autre ne l'est pas :

$$(\forall x : p(x) \vee q(x)) \Rightarrow (\forall x : p(x) \vee q(x)) \qquad \equiv Valide$$

$$(\forall x : p(x) \vee q(x)) \Rightarrow (\forall x : p(x)) \vee (\forall x : q(x)))$$
$$\equiv (\forall x : q(x)) \vee \neg q(x) \vee p(x) \qquad \equiv Contingente$$

Une procédure de semi décision existe certainement qui consiste à passer en revue toutes les preuves dans l'ordre canonique. Une preuve est un enchaînement d'invocations d'axiomes et de règles de production qui se termine par l'énoncé prouvé. Si la proposition donnée est prouvable il est impossible qu'on ne le remarque pas.

Certains fragments de la logique du premier ordre sont décidables et se situent en zone 3 du diagramme C-D. C'est évidemment le cas de la logique propositionnelle mais c'est aussi vrai de la logique des prédicats du premier ordre privée de fonctions ou encore de la logique monadique. On nomme ainsi ce cas particulier qui ne considère qu'une seule variable par prédicat donc p(x). L'étude des syllogismes rentre dans ce cas. Par contre, la logique dyadique est indécidable dès que le prédicat en question n'est pas l'égalité

L'une des grandes forces de la logique du premier ordre est de permettre d'énoncer un principe d'induction qu'on peut comprendre comme une extension infinie du Modus Ponens, $(p \wedge (p \Rightarrow q)) \Rightarrow q$. Portant sur une fonction, R, appliquée à la variable, x, il s'énonce :

$$(R(0) \wedge \forall x(R(x) \Rightarrow R(x+1))) \Rightarrow \forall x R(x)$$

Seuls les logiciens professionnels étudient les règles du raisonnement valide pour ce qu'elles sont. Faire des mathématiques, c'est greffer un ensemble d'axiomes supplémentaires sur les axiomes de la logique du premier ordre afin d'en enrichir le cadre.

Les logiques d'ordres supérieurs.

Nous avons vu dans le prologue que la logique d'ordre un n'est pas toute puissante et qu'en particulier elle n'est pas équipée pour formaliser des propositions du style,

"Le cardinal de l'ensemble, A, est fini"
"Le cardinal de l'ensemble, B, est $\aleph_0$"
"L'ensemble des sous-ensembles d'un ensemble dénombrable n'est pas dénombrable"…

Cela tient essentiellement au fait qu'une formule du premier ordre ne peut en aucun cas mélanger des objets de type "ensemble" avec des objets de type "ensemble d'ensembles". Il faut une logique d'ordre deux pour y parvenir. Dans une logique d'ordre deux, on exprime que l'ensemble, A, est fini en affirmant que toute fonction injective, f(x), de A sur lui-même est aussi surjective. De même pour exprimer la dénombrabilité de B, on affirme que deux sous-ensembles quelconques, b_1 et b_2, de B peuvent être mis en bijection. Dans ces formulations, c'est le mélange des quantifications, $\forall x$ et $\forall f$ (ou $\forall b$ et $\forall B$), qui est typique de l'ordre deux.

L'ennui c'est que la logique d'ordre deux cesse d'être sémantiquement complète d'où la difficulté d'y développer une bonne notion de preuve. Etant donné que les mathématiciens n'ont jamais souhaité se passer de celle-ci ni de pouvoir manipuler les différentes formes d'infini, il leur a fallu trouver une échappatoire. Cette échappatoire, c'est la théorie des ensembles : expressible dans un langage du premier ordre elle introduit ses propres axiomes relatifs à l'infini. Les théories développées dans ce cadre cessent d'être élémentaires, on les dit "étendues". Les logiques d'ordres supérieurs à un ne sont ni syntaxiquement complètes ni décidables ce qui les situe en zone 2 dans le diagramme C-D.

Deux systèmes situés en zone 4 : L'algèbre et la géométrie élémentaires.

Les systèmes qui occupent la zone 4 du diagramme C-D sont particulièrement agréables puisqu'ils sont syntaxiquement complets et décidables. A ce titre, ils incarnent l'idéal Hilbertien. Les exemples les plus connus sont la géométrie et l'algèbre élémentaires telles qu'axiomatisées par Tarski. On ne sépare habituellement pas ces deux systèmes car la procédure de décision qui est valable pour l'un est immédiatement transposable à l'autre.

Par définition, une théorie élémentaire s'exprime dans un langage du premier ordre et elle s'interdit de poser tout axiome relatif à la notion d'infini. De ce fait, l'algèbre et la géométrie élémentaires ne coïncident pas exactement avec ce que l'on étudie habituellement au lycée. Voici quelques points de repère à ce sujet.

L'algèbre élémentaire des corps réels clos est axiomatisable, en deux temps, sur les bases suivantes, dues à Tarski. D'abord un ensemble d'axiomes qui caractérisent tout corps additif et multiplicatif :

$$a+(b+c)=(a+b)+c \qquad a+0=a \qquad a+b=b+a \qquad a+(-a)=0$$
$$a\times(b\times c)=(a\times b)\times c \qquad a\times 1=a \qquad a\times b=b\times a \qquad a\neq 0\Rightarrow a\times a^{-1}=1$$
$$a>b\Rightarrow a\neq b \qquad ((a>b)\wedge(b>c))\Rightarrow a>c \qquad ((a>b)\wedge(c>0))\Rightarrow a\times c>b\times c$$
$$(a>b)\vee(a=b)\vee(a<b) \qquad (a>b)\Rightarrow a+c>b+c \qquad a\times(b+c)=(a\times b)+(a\times c)$$
$$1>0$$

Le système qui comprend les constantes, 0, 1, a, b, c, …, les variables, x, y, z, …, les opérateurs, +, -, x, et > , ainsi que les quantificateurs existentiel et universel, ne permet de construire que des fonctions polynomiales. Toute notion d'infinitude en est absente. On remarque que ces axiomes seraient, par exemple, satisfaits par l'ensemble des rationnels mais ce n'est pas ce que nous avons en vue. On étend le domaine des constantes et des variables aux nombres complexes en y ajoutant l'axiome suivant, dit "de fermeture" (également appelé "de continuité" car il équivaut au théorème des valeurs intermédiaires) :

$$\exists x : x^n + y_1 x^{n-1} + \cdots + y_n = 0$$

Cet axiome est en fait un schéma d'axiomes, un axiome par valeur de n. L'appellation "de fermeture" ou de façon équivalente la dénomination de "corps clos", évoque le fait que si, dans un premier temps, les nombres complexes sont les racines de n'importe quel polynôme à coefficients entiers, il sont aussi les racines de n'importe quel polynôme à coefficients complexes. Autrement dit, la notion de nombre complexe se suffit à elle-même pour l'étude des (systèmes d') équations polynomiales à un nombre quelconque mais fixé de variables.

On se restreint au corps des réels, en imposant à n d'être impair dans l'énoncé précédent et ajoutant l'axiome,

$$x\geq 0\Rightarrow(\exists y : yy=x)$$

Au premier ordre, les quantificateurs ne peuvent porter que sur les variables dès lors, les propositions suivantes sont parfaitement acceptables en algèbre élémentaire, peu importe qu'elles soient valides ou non :

"Tout polynôme de degré 1 possède une racine" : $\forall a \forall b \exists x : \neg(a = 0) \wedge (ax + b = 0)$

"Tout polynôme de degré 2 possède une racine" : $\forall a \forall b \forall c \exists x : \neg(a = 0) \wedge (ax^2 + bx + c = 0)$

etc …

Par contre la proposition suivante n'est pas acceptable dans le cadre élémentaire car elle exige une logique d'ordre deux où les quantificateurs agissent sur les fonctions :

"Tout polynôme, quel que soit son degré, possède une racine" : $\forall P \exists x : P(x) = 0$.

Tarski a démontré que l'algèbre élémentaire est complète et décidable et, mieux encore, a découvert une procédure effective qui décide toutes les propositions algébriques élémentaires. Cette procédure, très inefficace à l'origine, a été progressivement améliorée et elle est actuellement implémentée dans tous les logiciels évolués de calculs formels.

Décidabilité de l'algèbre élémentaire des réels.

L'existence d'une procédure de décision pour l'algèbre élémentaire des réels repose sur le fait qu'il y est toujours possible d'éliminer les quantificateurs d'une proposition bien formée. Par exemple, on a l'équivalence suivante :

$$\exists x : ax^2 + bx + c = 0 \Leftrightarrow (a = 0 \wedge b = 0 \wedge c = 0) \vee (a = 0 \wedge b \neq 0) \vee (a \neq 0 \wedge b \neq 0 \wedge b^2 - 4ac \geq 0)$$

Cet exemple peut être généralisé aux systèmes de polynômes de degrés quelconques. L'élimination des quantificateurs a pour conséquence de ramener un problème de logique du premier ordre à un problème de logique propositionnelle. Comme cette dernière est décidable, il s'en suit que l'algèbre des réels l'est également.

Concrètement, la procédure de décision équivaut à se demander si un système d'(in)équations possède ou non des solutions réelles dans un domaine imposé. Une généralisation d'un théorème dû à Sturm est capable de régler ce genre de questions au terme d'une stratégie finie. La procédure exacte, dont le détail fastidieux ne nous intéresse pas, est implémentée dans les logiciels de calcul formel, tel Mathematica qui résout le problème à l'aide de l'instruction, Reduce[]. Dans un premier temps elle élimine les quantificateurs puis elle passe à la forme normale disjonctive, $\bigvee_{i=1}^{m} \left(\bigwedge_{j=1}^{n} Q_{i,j} \right)$, et enfin elle traduit cette forme en un système algébrique auquel elle applique une variante du théorème de Sturm. La voici en action dans deux exemples proches. Dans un premier cas, la procédure répond correctement qu'aucune solution réelle n'existe :

$$\exists x, y : (x^2 + y^2 < 2) \wedge (x^3 + y^3 > 3)$$

```
Reduce[{x^2+y^2<2,x^3+y^3>3},{x,y}]
```
False,

tandis que dans ce cas voisin, elle trouve un ensemble non vide de possibilités :

$\exists x,y : (x^2 + y^2 < 2) \wedge (x^3 + y^3 > 2)$

```
Reduce[{x^2+y^2<2,x^3+y^3>2},{x,y}]
```

$$\left(\text{Root}[-2-4\,\#1+2\,\#1^3+\#1^4\,\&,\,1] < x < 1 \;\&\&\; \text{Root}[-2+x^3+\#1^3\,\&,\,1] < y < \sqrt{2-x^2}\right) \;||$$
$$\left(1 < x \le \text{Root}[-2-4\,\#1+2\,\#1^3+\#1^4\,\&,\,2] \;\&\&\; \text{Root}[-2+x^3+\#1^3\,\&,\,1] < y < \sqrt{2-x^2}\right) \;||$$
$$\left(\text{Root}[-2-4\,\#1+2\,\#1^3+\#1^4\,\&,\,2] < x < \sqrt{2} \;\&\&\; -\sqrt{2-x^2} < y < \sqrt{2-x^2}\right)$$

La procédure affiche les plages des valeurs des variables qui permettent de satisfaire la requête posée. On peut calculer numériquement les bornes des intervalles avec une précision aussi grande que l'on veut, par exemple, 20 chiffres significatifs :

```
N[Root[-2-4 #1+2 #1^3+#1^4&,1],20]
-0.56457945531766095218
```

Bien qu'il soit possible d'invoquer n'importe quel entier dans le système élémentaire de Tarski, sous la forme, 3=1+1+1, il n'est pas possible d'invoquer, en toute généralité, la notion d'entier. Par exemple, la proposition suivante ne fait pas partie du système formel de Tarski :

"L'équation, $x^3 + y^3 = z^3$, ne possède pas de solutions *entières* sur les variables, x, y, z".

En général, les équations diophantiennes ne sont, de ce fait, pas formalisables dans le système de Tarski et nous verrons le moment venu qu'il n'existe effectivement pas de procédure capable de décider l'existence de solutions entières.

Décidabilité de la géométrie élémentaire.

La géométrie élémentaire ne coïncide pas non plus avec l'idée qu'on pourrait s'en faire en se remémorant les leçons de géométrie apprises à l'école secondaire. Par géométrie élémentaire (synonyme : de Tarski), on entend cette partie de la géométrie qui est immédiatement traduisible analytiquement dans le cadre de l'algèbre du même nom grâce à un système de coordonnées adéquat. On ne s'étonnera pas d'apprendre que la décidabilité de l'algèbre élémentaire des réels algébriques s'étend immédiatement à la géométrie. La procédure de décision porte le nom de géométrie analytique élémentaire.

L'axiomatisation de la géométrie élémentaire a été faite par Tarski. On en trouvera le détail dans l'ouvrage de référence de Wolfram et il suffit d'en évoquer quelques uns. Désignant les points par des lettres minuscules, on pose par exemple :

$xy \equiv zz \Rightarrow x = y$ "Si le segment xy est congru à un segment qui se termine là où il commence c'est que les points x et y sont confondus".

$B(x,y,x) \Rightarrow x = y$ "Aucun point n'existe entre x et x. Les points sont donc indivisibles".

$(B(x,u,z) \wedge B(y,v,z) \Rightarrow \exists a(B(u,a,y) \wedge B(v,a,x)$ "Deux segments de droite joignant chacun un sommet différent d'un triangle au côté opposé se coupent à l'intérieur du triangle".

Ces axiomes sont, en gros, l'équivalent géométrique des axiomes algébriques qui caractérisent tout corps additif et multiplicatif. Ils ne suffisent cependant pas à développer une géométrie intéressante. Par exemple, dans cette géométrie restreinte, il n'est même pas possible de construire un triangle équilatéral sur une base connue. La méthode classique consistant à se servir d'un compas pour reporter la mesure de la base à partir de chacune des extrémités n'est, en effet pas certaine d'aboutir car rien dans les axiomes de base ne garantit que les cercles se coupent en un point. Il manque un axiome de continuité, en fait un schéma d'axiomes, que l'on note habituellement :

$$\exists a \forall x \forall y (\phi(x) \wedge \psi(y) \Rightarrow B(a,x,y) \Rightarrow \exists b \forall x \forall y (\phi(x) \wedge \psi(y) \Rightarrow B(x,b,y)$$

La géométrie élémentaire incorpore donc cet axiome qui joue, en géométrie, le rôle que l'axiome de fermeture jouait en algèbre.

La notion générique d'entier ne faisant pas partie de la théorie élémentaire, toute proposition invoquant des polygones à un nombre quelconque de côtés est exclue. Dès lors, le théorème relatif à la somme des angles d'un polygone convexe quelconque n'y est pas exprimable. Par contre, il est parfaitement possible de considérer n'importe quel polygone dont le nombre de côtés est spécifié, un triangle ou un pentagone, par exemple.

Evidemment, les passages à la limite tels ceux que l'on rencontre lorsqu'on calcule l'aire du cercle par la méthode des polygones inscrits et circonscrits d'Archimède ne font pas partie de la géométrie élémentaire. C'est à nouveau l'absence de statut de l'infini qui en est responsable. En revanche, rien n'interdit de s'intéresser à des problèmes géométriques compliqués qui concernent des courbes algébriques de degrés arbitrairement grands mais fixés, coniques (degré 2), ovales (degré 4), etc.

La relation entre l'algèbre des réels et la géométrie élémentaire peut se faire de la manière suivante. Cette géométrie parle de points, de droites, de cercles, etc. Elle pose, par exemple, à leur sujet des questions du genre :

- trois points sont-ils alignés ?
- trois droites sont-elles concourantes ?
- tel angle est-il égal à tel autre ?
- tels cercles se coupent-ils ? etc

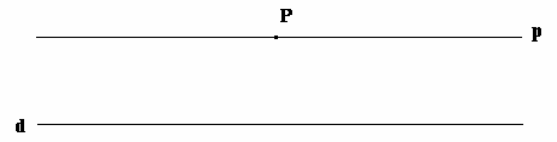
On répond à ces questions en codant :

- la position d'un point par ses coordonnées réelles,
- toute droite par deux points distincts,
- tout angle par les coordonnées de trois points qui le définissent dans l'ordre
- tout cercle par son centre et un point de son périmètre,
- etc pour les courbes de degrés supérieurs.

Toute question se règle alors par un système d'(in)équations polynomiales que la méthode algébrique précédente décide à coup sûr.

Un fragment de la géométrie élémentaire : La géométrie absolue.

Parmi les axiomes de la géométrie élémentaire, il en est un qui jouit d'un statut particulier, c'est l'axiome dit des parallèles. Euclide ne l'utilise, dans son premier livre, qu'à partir de la trentième proposition. Il pose que par un point, P, extérieur à une droite, d, ne passe qu'une seule parallèle, p, à cette droite.

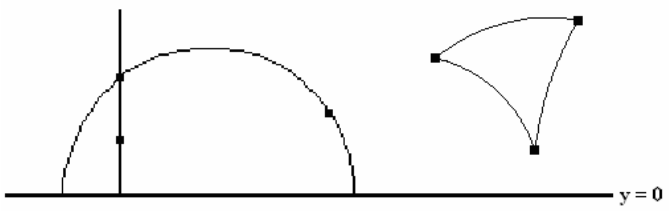


On appelle « géométrie absolue » le fragment de la géométrie élémentaire qui est amputé de ce postulat. Evidemment, suite à cette amputation, la géométrie absolue est syntaxiquement incomplète : la proposition relative aux parallèles devient indécidable et elle entraîne dans l'indécidabilité toutes les propositions qui ne peuvent se passer d'elles pour recevoir une preuve. Par exemple, en géométrie absolue, la somme des angles d'un triangle est inférieure ou égale à 180°. La proposition, "La somme des angles d'un triangle est égale à 180°" n'est donc ni prouvable ni réfutable en géométrie absolue, elle y est indécidable.

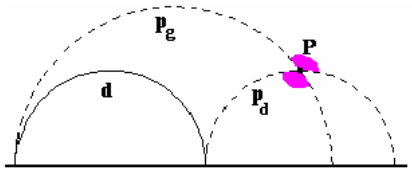
La géométrie absolue apparaît comme le socle potentiellement commun à deux géométries distinctes, l'une, euclidienne, qui accepte l'axiome des parallèles et l'autre, hyperbolique, qui le refuse. Inversement, on peut considérer ces deux géométries comme deux extensions distinctes de la géométrie absolue.

En géométrie plane hyperbolique, on pose que, par un point, P, extérieur à une droite, d, passent une infinité de parallèles à cette droite. Il ne faut pas prendre les mots au pied de la lettre : comme dans tout système formel, les vocables, « points» et « droites», ne sont que des symboles alphabétiques sans signification concrète. Un ordinateur que l'on programmerait pour étudier cette géométrie à partir des axiomes de base n'aurait pas de préjugés culturels concernant l'interprétation à donner à ces notions. Pour lui, les vocables « pommes » et « poires » feraient tout aussi bien l'affaire. Cependant, en pratique il est vrai que nous avons généralement en vue une interprétation particulière qui nous convient.

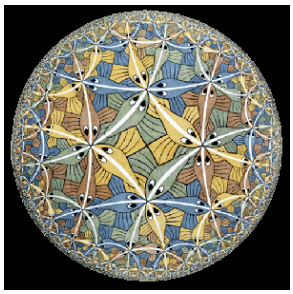
On peut proposer plusieurs interprétations isomorphes de la géométrie plane hyperbolique. La première, due à Poincaré, interprète l'espace de référence comme un demi-plan, ouvert, y>0. Les « points » restent des points et les « droites » sont les perpendiculaires à la frontière, y = 0, ou les demi-cercles centrés sur elle. On peut vérifier que les axiomes sont satisfaits : on a par exemple que par deux « points » passe une seule « droite » et que deux « droites » sécantes définissent un seul « point », etc.



On vérifie, sur la figure suivante, que le modèle se conforme à la version modifiée suivante de l'axiome des parallèles : par tout « point », P, extérieur à une « droite », d, passe une infinité de parallèles toutes comprises entre deux parallèles extrêmes, la « dernière » parallèle à gauche, p_g, et la « dernière » parallèle à droite, p_d (zones roses sur la figures).



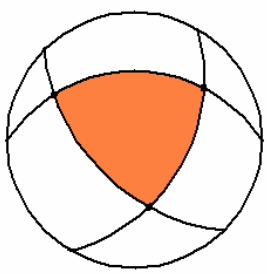
Une autre interprétation isomorphe assimile l'espace à un cercle ouvert, Γ, dont les droites sont tous les arcs de cercle centrés sur le périmètre. Le peintre Escher a basé l'un de ses tableaux sur une représentation de ce modèle.



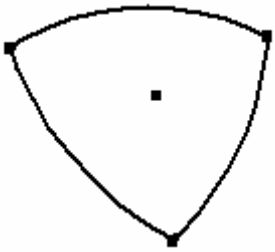
La géométrie elliptique.

La géométrie elliptique adopte une troisième version de l'axiome des parallèles, à savoir que, " Par un point extérieur à une droite, ne passe aucune parallèle à celle-ci". Rigoureusement parlant, la géométrie elliptique n'est pas une extension de la géométrie absolue : certains axiomes, peu nombreux il est vrai et que nous ne détaillons pas, ne sont pas communs aux deux systèmes. D'ailleurs on prouve, en géométrie elliptique, que la somme des angles d'un triangle est supérieure à deux droits, ce qui exclut d'emblée qu'il puisse s'agir d'une extension de la géométrie absolue.

La géométrie sur une surface sphérique est un modèle possible de cette géométrie elliptique où les points sont les couples de pôles opposés et les droites les grands cercles nécessairement tous sécants. En projetant stéréographiquement cette sphère, on construit un modèle isomorphe qui réutilise le cercle ouvert, Γ : les « droites » sont, cette fois, représentées par les arcs de cercle qui coupent Γ en deux points diamétralement opposés. La figure suivante représente un triangle dans cette géométrie.



Sur le papier, on peut concevoir autant de géométries que l'on veut, il suffit de modifier l'un ou l'autre axiome en priant pour que le système ne devienne pas contradictoire. Aucune géométrie n'est a priori « meilleure » ou « plus vraie » qu'une autre. Evidemment, le physicien a une conception utilitaire de la géométrie qui lui fait préférer le modèle particulier qui colle au mieux à la réalité. Un physicien ne se contente pas d'une définition axiomatique de la droite, il a naturellement tendance à chercher ce que pourrait être son incarnation dans le monde sensible. On pourrait penser à une corde que l'on tend entre deux points ou, mieux, à un faisceau laser qui relie la source au détecteur. On pense que dans le vide sidéral, loin de toute matière, trois droites ainsi définies traceraient un triangle dont la somme des angles vaudrait effectivement 180° et que la géométrie d'Euclide y serait d'application. Mais il suffirait de considérer le même triangle englobant un astre de masse importante pour que cette somme excède 180°. En présence de matière, la géométrie d'Euclide ne convient plus et on la remplace habituellement par une variante de type elliptique.



La suite de l'inventaire des systèmes formels usuels nous entraîne en zone 1 dans le diagramme, C-D. On y trouve, sans surprise, la toute puissante théorie des ensembles ainsi que toutes les théories qui lui empruntent un axiome relatif à l'infini. Etonnamment, on y trouve également les arithmétiques qui présentent un degré minimum de sophistication.

En zone 1 : La géométrie étendue d'Hilbert.

On aura remarqué, dans les sections précédentes consacrées à la géométie élémentaire, qu'on a évité de parler du système axiomatique d'Euclide car le géomètre grec n'a pas vraiment fait l'inventaire complet des axiomes qu'il utilise de façons souterraines. Les cinq malheureux axiomes qu'il a explicités ne suffisent pas pour programmer une machine qui ferait le travail de déduction logique à la place de l'homme.

La géométrie élémentaire de Tarski n'est qu'un fragment d'une géométrie étendue qui s'exprime dans le langage plus puissant de la théorie des ensembles. Cette extension lui confère un pouvoir d'expression beaucoup plus grand qui incorpore les notions d'infinis. Elle peut formuler des théorèmes concernant des polygones à un nombre quelconque de côtés ou calculer des longueurs de courbes, des aires ou des volumes.

Ce pouvoir étendu a un prix qui est la perte de décidabilité. Certes, un grand nombre de schémas de problèmes géométriques restent heureusement décidables mais il n'existe plus de procédure mécanique effective qui décide à coup sûr toutes les instances de la géométrie étendue.

La traduction analytique de la géométrie étendue est possible mais elle nécessite le recours aux ressources de l'analyse. On y parvient habituellement en munissant l'espace d'une métrique adaptée. Le modèle euclidien opte pour la métrique,

$ds^2 = dx^2 + dy^2$,

tandis que le modèle hyperbolique du demi-plan ouvert opte pour la métrique,

$$ds^2 = \frac{dx^2 + dy^2}{y^2}$$

Cette métrique garantit en particulier que toute "droite" de ce modèle respecte l'axiome qui exige qu'elle puisse être prolongée indéfiniment dans ses deux directions, un fait qui ne saute pas aux yeux lorsqu'on regarde les figures trompeuses, en "demi-cercles".

Le « segment de droite » qui joint deux points, (x_1, y_1) et (x_2, y_2), est encore défini comme étant le plus court chemin entre ces points. Vu la métrique adoptée, cela revient à minimiser l'intégrale suivante,

$$\int_1^2 \frac{1}{y}\sqrt{1+\dot{y}^2}\,dx = \int_1^2 f(y,\dot{y},x)dx$$

La solution de ce genre de problème passe par la résolution de l'équation d'Euler associée,

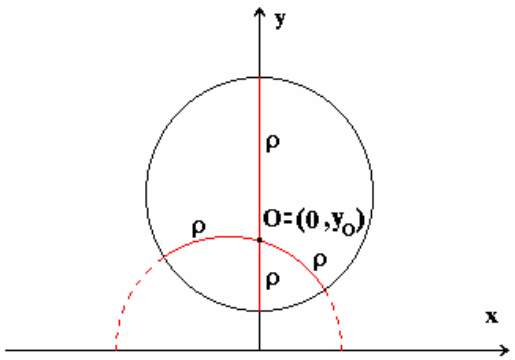
$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = 0 \qquad \Rightarrow \qquad (x - x_C)^2 + y^2 = R^2$$

dont la représentation dans le demi-plan est bien un arc de demi-cercle centré sur la frontière et passant par les deux points. « L'abscisse du centre » et « le rayon » valent respectivement :

$$x_C = \frac{x_1^2 - x_2^2 + y_1^2 - y_2^2}{2(x_1 - x_2)}$$

$$R = \frac{1}{2|x_1 - x_2|}\sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2][(x_1 - x_2)^2 + (y_1 + y_2)^2]}$$

Curieux sont les « cercles » dont la définition n'a évidemment pas changé, à savoir le lieu des points équidistants d'un centre fixe. Ils ont l'aspect d'un cercle euclidien mais ils sont décentrés afin de tenir compte de la métrique particulière à cette géométrie.



Le « cercle » « centré » sur O et de rayon, ρ, a pour équation, $x^2 + y^2 - 2yy_0 ch\rho + y_0^2 = 0$, et on peut vérifier que dans la métrique imposée, chaque « rayon » émanant de O (en trait plein rouge sur la figure) possède la longueur constante, ρ. Ce calcul n'est élémentaire que sur les deux « rayons » rectilignes :

$$\int_{y_0}^{y_0 e^{\rho}} \frac{dy}{y} = \int_{y_0 e^{-\rho}}^{y_0} \frac{dy}{y} = \rho \,.$$

De même, la longueur de la circonférence vaut :

$$2\int_{y_0 e^{-\rho}}^{y_0 e^{\rho}} \frac{dy}{y}\sqrt{1+(dx/dy)^2} = 2\pi \sinh(\rho),$$

soit davantage que dans la version euclidienne. Enfin, la surface du cercle est donnée par l'intégrale :

$$\iint \frac{dxdy}{y^2} = 2\pi(ch\rho - 1).$$

Par l'usage qu'elle fait de l'analyse infinitésimale, la géométrie différentielle se range obligatoirement dans la zone 1 du diagramme C-D.

La métrique associée au modèle hyperbolique à cercle ouvert suppose une métrique un peu différente,

$$ds^2 = 4\frac{dx^2 + dy^2}{(1 - x^2 - y^2)^2}$$

mais un simple changement de variables renvoie au modèle du demi-plan de Poincaré, signe qu'il s'agit bien d'interprétations isomorphes.

Par l'usage qu'elle fait de la notion de continu, la géométrie différentielle se situe en zone 1, dans le diagramme, C-D. Dit autrement, elle s'exprime dans le langage de la théorie des ensembles.

Les avatars de la théorie des ensembles.

Les physiciens ayant largement adhéré au système formel ZF(C) de la théorie des ensembles, il est bon qu'ils sachent à quoi ils se sont engagés.

La théorie naïve des ensembles, basée sur les axiomes d'extension et de compréhension,

$X = Y \Leftrightarrow \forall z(z \in X \Leftrightarrow z \in Y)$ (axiome d'extension : les éléments sont cités)

$\exists X \forall z(z \in X \Leftrightarrow P(z))$ (axiome de compréhension : la propriété d'appartenance est fixée)

est contradictoire, ainsi que le révèle le paradoxe de Berry : « Soit l'ensemble, A, des entiers caractérisables complètement par une phrase rédigée en français correct et comportant moins de 100 caractères prélevés dans l'alphabet {a,b, …, z, _} ».

On peut, au maximum, écrire 27^{100} chaînes de caractères dans cet alphabet et la plupart d'entre elles ne caractérisent pas un entier. A possède donc (beaucoup) moins que 27^{100} éléments et son complémentaire est largement non vide. Ce complémentaire possède nécessairement un plus petit élément, x. Mais « x est le plus petit entier non définissable par une phrase française de moins de cent caractères » ne comporte que 96 caractères : c'est une définition parfaitement valable dans le système de Cantor d'un nombre qui devrait simultanément appartenir à A et à son complémentaire, une contradiction.

Comment sortir de l'ornière ? Une solution radicale serait d'abandonner l'axiome de compréhension pour ne garder que l'axiome d'extension. C'est en gros le point de vue constructiviste de l'informatique théorique. L'immense majorité des mathématiciens a refusé d'épouser ce point de vue et a cherché une autre voie. L'issue trouvée par Frege a consisté, dans un premier temps, à restreindre la validité du principe de compréhension en interdisant les définitions qui se servent de la rhétorique. Plus formellement, les axiomes de compréhension qui définissent de « bons » ensembles doivent respecter la syntaxe de la logique dite du premier ordre, ce que l'énoncé de Berry ne fait pas :

$\forall a_1,\cdots,a_n : \tau \quad \exists A : Ens_\tau \quad \forall x : \tau \quad (x \in A \Leftrightarrow F(x, a_1,\cdots,a_n))$

où les notations, :τ, et, :Ens$_\tau$, se réfèrent à l'appartenance à un type donné (entier, réel, ensemble d'entiers, …), on aurait pu écrire de façon équivalente, $\forall a_1,\cdots,a_n \in N$.

Pensant avoir réparé la théorie des ensembles à peu de frais, Frege reçut un choc lorsque, à deux doigts de publier sa méthode, Russell lui fit remarquer qu'elle n'échappait pas à une autre contradiction qui surgit dès qu'on considère l'ensemble de tous les ensembles.

« L'ensemble de tous les ensembles qui ne font pas partie d'eux-mêmes » est une définition d'ensemble parfaitement valide dans le système de Frege, puisqu'il se note : $\{ X : Ens \quad X \notin X \}$. Cependant son existence même est contradictoire, vu qu'on aurait simultanément, $X \in X$ et $X \notin X$.

Appelons "ordinaire" un ensemble qui n'appartient pas à lui-même. Les ensembles usuels, sont de ce type, par exemple, l'ensemble des pommes n'est pas une pomme, il n'appartient donc pas à lui-même. A l'opposé, appelons « extraordinaire » un ensemble qui appartient à lui-même. On peut se demander s'il existe de

tels ensembles. En cherchant un peu, on trouve que l'ensemble des ensembles est effectivement extraordinaire. Considérons, à présent, l'ensemble des ensembles ordinaires. Est-il ordinaire ou extraordinaire ? Il ne peut pas être ordinaire sinon il appartiendrait à l'ensemble des ensembles ordinaires, c'est-à-dire à lui-même et par conséquent, il serait extraordinaire. Il ne peut pas davantage être extraordinaire sinon, par définition des ensembles extraordinaires, il devrait appartenir à lui-même donc être ordinaire. En conclusion, l'ensemble des ensembles ordinaires est contradictoire : il ne peut être ni ordinaire ni extraordinaire. Il n'est pas difficile de se convaincre que c'est l'auto référence qui crée un problème dans la définition des ensembles ordinaires et extraordinaires.

De toute évidence, la formulation de Frege demeure contradictoire et une réparation supplémentaire s'impose. L'histoire a dégagé deux issues également valables, dont une a finalement été préférée pour des raisons de commodité.

La première solution, due à Russell, n'a pas survécu à son auteur. Elle consistait à n'autoriser que les relations d'appartenance entre objets d'un type, τ, et ensembles d'objets de ce type, eux-mêmes de type, Ens_τ. Dans l'optique de Russell, il y aurait lieu de ne considérer que des objets de base, puis des ensembles d'objets de base, puis des ensembles d'ensembles des objets de base, etc, les relations d'appartenance ne prenant sens qu'entre objets de classes contiguës. Dans ce système, les relations auto référentielles du type, $X \in X$ et $X \notin X$ sont interdites et le paradoxe disparaît. Cette théorie est parfaitement défendable mais elle mène à des complications inextricables qui ont découragé la plupart des mathématiciens. Il suffit de contempler une seule page des Principes édités par Russell et Whitehead pour comprendre la colère de Poincaré qui trouvait ridicule d'avoir besoin de trois pages pour démontrer que un et un font deux.

L'autre solution, effectivement retenue, consiste à restreindre davantage le principe de compréhension en remplaçant l'exigence syntaxique,

$$\forall a_1,\cdots,a_n : \tau \;\; \exists A : Ens_\tau \;\; \forall x : \tau \;\; (x \in A \Leftrightarrow F(x,a_1,\cdots,a_n)),$$

par la suivante :

$$\forall a_1,\cdots,a_n : \tau \;\; \forall A : Ens_\tau \;\; \exists B : Ens_\tau \;\; \forall x : \tau \;\; (x \in B \Leftrightarrow (x \in A \wedge F(x,a_1,\cdots,a_n)))$$

ou, si l'on préfère une formulation plus conviviale,

$$\exists X \forall z (z \in X \Leftrightarrow ((z \in A) \wedge P(z)))$$ (axiome de séparation).

Il ne s'agit plus d'autoriser tous les ensembles, X, créés ex nihilo mais à séparer, à l'intérieur d'un ensemble préexistant, A, les éléments qui vérifient P. Pour cette raison, on appelle axiome de séparation cette version restreinte de l'axiome de compréhension.

Pour un type donné, τ, deux situations sont alors possibles. Soit il existe un ensemble, Ω_τ, de tous les objets de type, τ, et alors compréhension et séparation mènent aux mêmes définitions ensemblistes soit un tel ensemble n'existe pas et c'est là le lieu de la restriction cherchée.

Le passage par le détour d'ensembles préexistants crée une complication qui ne se résout qu'au travers de l'adoption d'axiomes supplémentaires tels ceux, dits de la paire, de l'union et des parties, qui sont des axiomes de compréhension élémentaires à partir desquels les autres suivront. On espère que ces axiomes ne recèlent plus aucune mauvaise surprise paradoxale et que leur contenu suffit pour définir un univers suffisamment riche.

$\forall a,b : \tau \;\; \exists A : Ens_\tau (a \in A \wedge b \in A)$ (paire)

$\forall A : Ens_{Ens_\tau} \;\; \exists B : Ens_\tau \;\; \forall x : \tau (\exists X : Ens_\tau \;\; (x \in X \wedge x \in A) \Rightarrow x \in B)$ (union)

$\forall A : Ens_\tau \;\; \exists B : Ens_{Ens_\tau} \;\; \forall X : Ens_\tau (X \subseteq A \Rightarrow X \in B)$ (parties)

$\forall a_1,\cdots,a_n : \tau \;\; \forall A : Ens_\tau \;\; \exists B : Ens_\tau \;\; \forall x : \tau \;\; (x \in B \Leftrightarrow (x \in A \wedge F(x,a_1,\cdots,a_n)))$ (séparation)

L'axiome des parties rend légal l'ensemble des sous-ensembles d'un ensemble donné. Si l'ensemble donné est fini, rien de particulier n'en découle. Par contre si l'ensemble est celui des entiers, N, l'ensemble, non dénombrable de ses parties prend une existence légale. C'est une porte qui s'ouvre sur un infini non constructif.

La théorie provisoirement définitive des ensembles a été axiomatisée par Zermelo. Au fil du temps, ce système a subi quelques retouches par adoptions successives de :

- l'axiome de l'infini, soit informellement, $\exists a(\varnothing \in a \wedge \forall x \in a(succ(x) \in a))$, qui légitime les ensembles tels que N (Zermelo),

- l'axiome de fondation soit, informellement, $\forall a(a \neq \varnothing \Rightarrow \exists b \in a(b \cap a) = \varnothing))$, qui exprime que tout ensemble est pur, c'est-à-dire qu'il appartient obligatoirement à la fermeture récursive des parties et unions d'ensembles purs en partant de $\varnothing$ (Zermelo-Fraenkel (ZF)),

- l'axiome du choix (optionnel et controversé parce qu'inutilement puissant) qui autorise de construire un ensemble à partir d'un ensemble d'ensembles non vides en sélectionnant « simultanément » (et non plus récursivement) un élément dans chaque sous-ensemble même lorsque ceux-ci sont en nombre infini (ZFC). L'axiome du choix équivaut à postuler que tout ensemble peut être bien ordonné.

L'hypothèse du continu généralisée entretient des rapports étroits avec l'axiome du choix dans la théorie ZF. Sierpinski a montré en 1947 que l'hypothèse généralisée du continu a pour conséquence l'axiome du choix dans ZF. Par contre ZF seule n'implique pas l'axiome du choix, comme l'a montré Paul Cohen. Ceci signifie que l'hypothèse du continu est beaucoup plus forte que l'axiome du choix. Sachant que ce dernier est déjà contesté par nombre de mathématiciens, autant dire que personne n'est prêt à incorporer l'hypothèse du continu dans l'arsenal axiomatique de ZFC. Signalons encore que des travaux récents dus à Woodin laissent penser que l'hypothèse du continu pourrait être fausse dans une extension bien choisie de ZF.

L'axiome du choix est ineffectif de naissance et à ce titre il est refusé par tout mathématicien constructiviste : en effet, il ne se préoccupe absolument pas de la façon dont il faudrait s'y prendre pour sélectionner un élément simultanément dans une collection infinie d'ensembles : il se contente de poser que cela est possible en espérant que cela ne mènera jamais à une contradiction ultérieure.

L'axiome du choix entraîne toutes sortes de conséquences qui si elles ne sont pas contradictoires n'en sont pas moins nettement contraires à l'intuition, telles l'existence de sous-ensembles de R non mesurables pour la mesure de Lebesgue ou l'existence d'un bon ordre sur les réels assortis de l'impossibilité, dans les deux cas, de produire le moindre exemple. De même, il entraîne le paradoxe de Banach-Tarski sur la multiplication des sphères dans R_3. Tout physicien est en droit de se méfier de l'axiome du choix : le paradoxe de Banach-Tarski, en particulier, entraîne de telles dérives physiques que l'on doit y voir une raison suffisante de ne pas suivre les mathématiciens sur ce terrain. Cela dit, l'essentiel de l'analyse infinitésimale utile au physicien peut être formulée sans faire appel à cet axiome.

Le système ZF(C) n'a, à ce jour, connu aucun épisode intrinsèquement paradoxal. Rien n'exclut cependant que cela puisse se produire même si cela est regardé comme hautement improbable. Au cas où cela se reproduirait malgré tout, force serait de procéder aux réparations qui s'imposeraient. Il n'est pas question de poursuivre ici l'axiomatique de la théorie des ensembles qui est une profession en soi. Le lecteur intéressé pourra consulter le remarquable cours de P. Dehornoy, déjà cité en référence.

La théorie des ensembles ne s'est fixé aucune limite dans l'usage de la notion d'infini : elle occupe de ce fait la zone 1 du diagramme C-D.

Les arithmétiques de Robinson et de Peano.

Il est étonnant que l'arithmétique des entiers naturels occupe également la zone 1 du diagramme C-D. Ceci montre d'emblée la grande différence de complexité entre l'algèbre, la géométrie et l'arithmétique élémentaires.

L'arithmétique de Robinson est la formalisation la plus simple d'une arithmétique universelle dans le cadre d'un langage du premier ordre. Elle pose l'existence de '0' et de la fonction unaire, Sa, qui désigne le "successeur" de a. Elle définit, en sus, les opérations binaires, "addition" et "multiplication". Ses axiomes s'écrivent comme suit :

ax1 :	$0 \neq Sa$	0 n'est "le successeur" d'aucun "nombre".
ax2 :	$Sa = Sb \Rightarrow a = b$	"L'égalité" des "successeurs" implique celle des "nombres".
ax3 :	$a + 0 = a$	Définit "l'addition" conjointement avec l'axiome suivant.
ax4 :	$a + Sb = S(a + b)$	
ax5 :	$a \times 0 = 0$	Définit "la multiplication" conjointement avec l'axiome suivant.
ax6 :	$a \times Sb = a \times b + a$	
ax7 :	$a \neq 0 \Rightarrow \exists b : a = Sb$	Pose qu'il n'y a pas de limite à la taille d'un entier.

Les guillemets sont là comme d'habitude pour rappeler que les substantifs qu'ils encadrent n'ont à ce stade aucune signification particulière. Cependant, il existe une interprétation "naturelle" (synonyme : standard) où :

- $\underbrace{SS\cdots S}_{n\ fois}0$ représente l'entier naturel, n, que tout le monde connaît,

- les symboles littéraux, a et b, désignent l'un quelconque des entiers naturels,

- ces axiomes mènent à l'arithmétique élémentaire capable de prouver que, par exemple, 2+2=4.

On prouve cette proposition, sous la forme, $SS0 + SS0 = SSSS0$, en appelant les axiomes dans l'ordre suivant :

$$SS0 + SS0 \underset{ax4}{=} S(SS0 + S0) \underset{ax4}{=} SS(SS0 + 0) \underset{ax3}{=} SSSS0 .$$

On réfute une proposition, par exemple, $S0+S0=S0$, en mettant en évidence une contradiction au sein du système formel, ici un conflit avec l'axiome 1 :

$$S0+S0 \underset{ax4}{=} S(S0+0) \underset{ax3}{=} SS0 \underset{ax2}{\Rightarrow} S0=0 \quad ???$$

Ce système est certainement syntaxiquement incomplet car un grand nombre de propositions y sont indécidables, tel l'énoncé, $a+b=b+a$, que l'axiomatique de Robinson ne permet ni de démontrer ni de réfuter en toute généralité. Par contre, elle démontre séparément, comme suit, chaque instance de ce schéma, par exemple, $SS0+S0=S0+SS0$:

$$SS0+S0 \underset{ax4}{=} S(SS0+0) \underset{ax3}{=} SSS0 \underset{ax3}{=} SS(S0+0) \underset{ax4}{=} S(S0+S0) \underset{ax4}{=} S0+SS0$$

Autrement dit, le système de Robinson est capable de démontrer la commutativité de l'addition instance après instance mais il échoue à les condenser toutes en une seule formule. D'autres indécidables du système de Robinson sont :

$0+a=a$ ou $a+(b+c)=(a+b)+c$, …

La formulation suivante est équivalente mais plus savante : le système de Robinson est sujet à l'ω-incomplétude. On veut dire par là qu'il existe des chaînes de théorèmes tous indexés par un entier, i, telles que la formule quantifiée qui les condense toutes n'est pas un théorème.

Il n'est généralement pas question de modifier l'un quelconque des six premiers axiomes de Robinson car cela ferait perdre l'interprétation élémentaire de l'arithmétique apprise dès l'école primaire. On peut supprimer les axiomes relatifs à la multiplication (ou alternativement ceux relatifs à l'addition) mais les arithmétiques résultantes, dites de Pressburger, sont tellement rudimentaires qu'elles ne présentent aucune intérêt pratique. En particulier, elles sont complètes et décidables, ce qui les refoule en zone 4.

L'importance du système de Robinson est surtout théorique. Elle est due au fait que c'est le système connu le plus simple qui appartient à la zone 1. Ce n'est toutefois pas le système que les arithméticiens utilisent en pratique car ils souhaitent éradiquer l'indécidabilité de certaines propositions telle, a+b=b+a, qu'ils considèrent comme définitivement vraies et pour lesquelles ils exigent qu'une démonstration soit possible. Ils y sont parvenus en élargissant l'axiomatique tout en espérant ne pas l'avoir rendue contradictoire. Concrètement, Peano a étendu le système de Robinson en profitant de la possibilité qu'offre la logique du premier ordre de formaliser le principe de récurrence. Il a remplacé l'axiome 7 par le schéma d'axiomes suivant :

ax7' : $(p[0] \wedge \forall a : (p[a] \Rightarrow p[Sa])) \Rightarrow \forall b : p[b]$

On pourrait être tenté de l'écrire sous la forme synthétique,

ax 7" : $\forall p : (p[0] \wedge \forall a : (p[a] \Rightarrow p[Sa])) \Rightarrow \forall b : p[b]$,

qui présenterait l'avantage de condenser le principe en un axiome unique. Mais ce faisant, on sortirait de la logique du premier ordre puisqu'on mélangerait des quantifications portant sur des objets, p et b, de types différents. En adoptant la forme, 7', réitérée de façon récursivement énumérable pour toutes les propriétés, p , définissables par une formule, on reste au premier ordre au prix du remplacement d'un axiome unique par un schéma infini mais récursivement énumérable d'axiomes. Nous verrons dans un instant que l'extension de Robinson dans Peano n'a pas ce problème puisque l'axiome de récurrence est précisément ce qu'il faut pour formaliser cette condensation.

C'est l'arithmétique de Peano du premier ordre, qui sert de cadre de référence consensuel pour la théorie des nombres entiers. Bien que l'adoption de l'axiome d'induction ait permis de lever l'indécidabilité des énoncés élémentaires mentionnés, $0+a=a$, $a+(b+c)=(a+b)+c$, … , elle demeure syntaxiquement incomplète et indécidable. Nous verrons d'ailleurs sous peu que cette arithmétique est en fait essentiellement syntaxiquement incomplète.

Il est tout à fait remarquable que des systèmes formels aussi simples d'aspect, que les arithmétiques de Robinson et de Peano soient essentiellement syntaxiquement incomplets, à égalité avec la théorie des ensembles. Ils sont même universels au sens de Gödel. Présentée tout d'abord informellement, cette étrange propriété affirme la capacité qu'ont ces arithmétiques d'émuler n'importe quel système y compris eux-mêmes au travers d'un codage bien choisi.

L'émulation et l'universalité au sens de Gödel.

Nous venons d'utiliser deux mots, universalité et émulation, que nous avons déjà utilisés dans l'approche Turingienne de la calculabilité. L'émulation et l'universalité Gödelienne sont d'une nature différente quoique apparentée : elles concernent les pouvoirs d'expression des différents systèmes formels.

Nous avons vu que l'émulation au sens de Turing signifie qu'un système calculatoire est capable de "piloter" le calcul d'un autre système au travers d'un codage effectif adéquat. L'universalité au sens de Turing se réfère, quant à elle, au fait que certains systèmes calculatoires sont capables d'émuler tous les systèmes, y compris eux-mêmes.

Lorsqu'on utilise le terme, émulation, dans le contexte Gödelien, on fait allusion à la capacité qu'ont certains systèmes formels de "piloter" à distance les preuves qui sont déroulables dans un autre système. De même, l'universalité au sens de Gödel se réfère au fait que certains systèmes formels sont capables d'émuler les preuves déroulables dans n'importe quel système, y compris eux-mêmes, du plus élémentaire au plus sophistiqué. Il semble inutile de préciser systématiquement dans quel sens on utilise les mots, émulation et universalité, puisque le contexte, systèmes calculatoires ou systèmes formels, parle de lui-même.

On dit qu'un système formel, Σ_A, est capable d'émuler un système, Σ_B, s'il est capable d'énoncer dans son propre langage un résultat de prouvabilité (ou de réfutabilité) qui concerne un énoncé, q, de Σ_B. Plus précisément, quelle que soit q, formulable au sein de Σ_B , Σ_A doit pouvoir formuler, en son sein, les propositions suivantes :

Dans Σ_A, on formule, p_1 : "La proposition, q, est prouvable dans Σ_B "

Dans Σ_A, on formule, p_2 : "La proposition, q, est réfutable dans Σ_B ".

L'émulation Gödélienne est triviale si Σ_B n'est qu'un fragment de Σ_A. Par contre, le cas inverse est intéressant, qui voit un système émuler une de ses extensions ou d'ailleurs n'importe quel système plus puissant que lui. Les propositions p_1, p_2 sont éventuellement prouvables ou réfutables au sein de Σ_A, et on peut en dire autant de q au sein de Σ_B. Les cas

dignes d'intérêt sont ceux qui ne mènent à aucune contradiction entre les affirmations faites au sein de Σ_B et de Σ_A. Concrètement on peut avoir :

- Si q est prouvable dans Σ_B, alors p_1 (resp. p_2) est soit prouvable (resp. réfutable) soit indécidable dans Σ_A.

- Si q est réfutable dans Σ_B, alors p_1 (resp. p_2) est soit réfutable (resp. prouvable) soit indécidable dans Σ_A.

- Si q est indécidable dans Σ_B, alors p_1, p_2 sont indécidables dans Σ_A.

Dans le premier cas, par exemple, il suffit que p_1 soit prouvable pour qu'on ait réussi à prouver, au sein de Σ_A, qu'il existe une preuve de q dans Σ_B. On a alors réussi à démontrer que q est prouvable dans Σ_B sans sortir de Σ_A. Cette preuve indirecte est la raison pour laquelle on parle d'émulation. Il est essentiel de comprendre qu'on n'a nullement apporté la preuve que q est directement prouvable dans Σ_A et il est parfaitement possible que q demeure un indécidable de Σ_A. Par contre, il est exclu que q soit réfutable dans Σ_A sous peine de contradiction majeure. Enfin, si q est indécidable dans Σ_B, il l'est également dans Σ_A.

On connaît plusieurs systèmes universels au sens de Gödel, en particulier et sans surprise, la très puissante théorie des ensembles, ZF(C). Toutefois, nous verrons, lors de l'étude des théorèmes de Gödel, que tous les systèmes qui contiennent l'arithmétique de Robinson sont déjà universels dans ce sens, en particulier l'arithmétique habituelle de Peano.

Par exemple, nous verrons qu'il est possible d'exprimer dans l'arithmétique de Robinson que "Le grand théorème de Fermat est prouvable dans le cadre ZFC". Vu que nous savons que cette proposition est vraie dans ZFC, aucun contre exemple n'est à redouter et il en résulte que nous avons prouvé que Robinson prouve que ZFC prouve Fermat. Ceci ne constitue cependant pas une preuve directe de Fermat dans Robinson et il est parfaitement possible que Fermat soit un indécidable de Robinson auquel cas il existerait un modèle non standard de Robinson où Fermat serait faux.

Rapports entre l'universalité des systèmes calculatoires et celle des systèmes formels.

Si le but de tout système calculatoire est de calculer des fonctions, le but des systèmes formels est de démontrer des théorèmes dans un cadre axiomatique préétabli. Quel rapport peut-il exister entre ces deux types de systèmes? La réponse à cette question confronte deux conceptions différentes des mathématiques.

Les mathématiques sont généralement enseignées d'un point de vue axiomatique : on démontre des théorèmes à partir des axiomes de base sans vraiment se préoccuper de l'aspect calculatoire concomitant. Dans cette optique, un théorème d'existence qui affirme que tel problème possède au moins une solution est considéré comme un progrès notable même si la démonstration ne dit rien de la forme exacte de cette solution. L'autre point de vue est, on s'en doute, essentiellement constructif en ce qu'il recherche des algorithmes capables de révéler la forme exacte de la solution. Les deux points de vue sont cependant liés.

Nous allons montrer que toute procédure algorithmique peut être vue comme un système formel particulier. Plus formellement, nous allons montrer, qu'étant donnée une machine de Turing qui calcule une certaine fonction, il existe un système formel qui est capable d'en faire autant au travers d'un codage effectif des données et des résultats. Il en résultera que si la MT est universelle au sens de Turing, le système formel associé le sera lui aussi, au sens de Turing, dans un premier temps, puis au sens de Gödel.

La parenté entre les universalités aux sens de Turing et de Gödel peut être mise en évidence par l'étude des "Systèmes Multi Voies", en abrégé, SMV, souvent appelés "Systèmes Semi-Thue", du nom du mathématicien norvégien, Axel Thue.

Les SMV sont des systèmes formels d'un genre particulier qui reposent sur un axiome unique, qui tient lieu de théorème de la première génération et sur un ensemble de règles de productions. Les générations successives de théorèmes s'obtiennent en épuisant les règles de productions sur l'ensemble des théorèmes obtenus à la génération précédente.

Voici, pour fixer les idées, un SMV particulièrement simple dont l'alphabet ne comprend que deux symboles, A et B. Imaginons que l'axiome unique s'écrive, BABB, et que les règles de productions soient définies par les relations :

AB→AAB BA→ABB AABB→BAB

La suite des théorèmes, rangés dans l'ordre des générations successives, se calcule comme suit :

```
NestList[Union[Flatten[StringReplaceList[#,{"AB"→"AAB","BA"→"ABB","AABB"→"
BAB"}]]]&,{"BABB"},2]
```

```
{{BABB},{ABBBB,BAABB},{AABBBB,ABBABB,BAAABB,BBAB}, … }
```

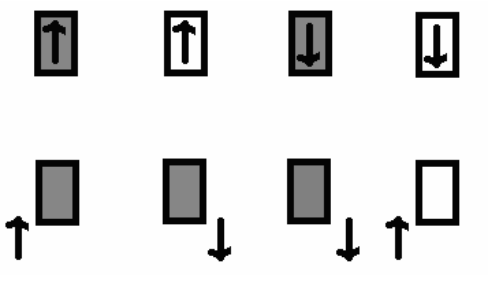
La connexion entre MT et SMV est réglée par le théorème d'émulation suivant :

Pour toute machine de Turing, MT, fonctionnant sur deux caractères,□=0 et ■=1 , il existe un SMV et une fonction d'encodage effectif, enc[x,y], tels que le SMV dérive la chaîne, enc[a,b], à partir d'une unique cellule noire, ■ , si et seulement si la MT imprime b sur la donnée a.

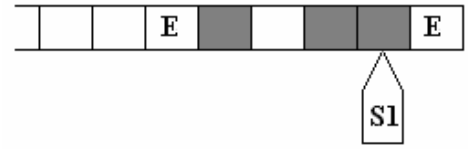
De manière informelle, ce théorème affirme qu'étant donnée une machine de Turing, il existe un SMV qui est capable de l'émuler au travers d'un codage effectif. On démontre ce théorème en indiquant la manière de construire effectivement la suite des instructions du SMV à partir des instructions de la machine de Turing ainsi qu'en fournissant son mode d'emploi.

Une machine de Turing portant sur s états et 2 caractères (ce qui, rappelons-le, n'est jamais une restriction) s'avère équivalente à un SMV dont l'alphabet possède, 8+s, caractères et, 2sK+2s+7, instructions. La construction effective suivante est due à Szudzik et nous l'illustrons sur l'exemple de la MT 1528 (= 010111111000_{bin}), à 2 états et 2 caractères, dont la table d'instructions est :

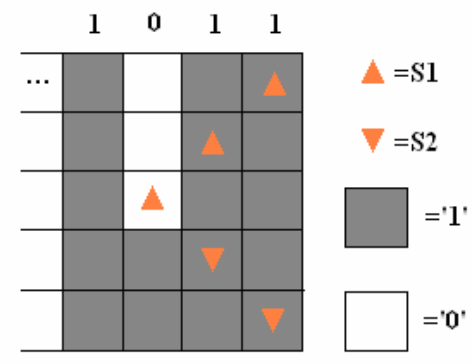
```
{1,1}→{1,1,-1}    {1,0}→{2,1,1}    {2,1}→{2,1,1}    {2,0}→{1,0,-1}
```



Afin de rendre les choses plus claires, nous suivrons l'évolution de la MT sur la donnée, '1011'=■•■■, qu'on inscrit initialement de gauche à droite sur sa bande semi infinie. La tête de lecture, initialement dans l'état, S1, pointe sur la cellule autorisée la plus à droite.



Le délimiteur, E, placé à l'extrême droite de la bande fixe la zone interdite d'accès à la tête de lecture. Si la MT tente d'y accéder sous prétexte que la table d'instruction le lui demande, elle s'arrête automatiquement laissant la bande afficher tel quel le résultat de son calcul. Dans l'exemple choisi, cela se produit lorsque la MT tente une cinquième instruction. Au bilan, la MT qui est partie de la donnée, w = ■•■■, affiche le résultat, ■■■■.



Le SMV associé utilise, 8+s=10, caractères, {■, C, S, 0, 1, 2, R, L, E, ?}, un axiome unique, ■, et, 2sK+2s+7=19, instructions réparties en six groupes qui traduisent mot à mot les actions distinctes effectuées par la machine de Turing :

```
(1) : "■"→"E?","?"→"C0?C0","?"→"C1?C1","?"→"S1E"
(2) : "C1S1"→"C1L1","C0S1"→"C1R2","C1S2"→"C1R2","C0S2"→"C0L1"
(3) : "C0L1"→"S1C0","C0L2"→"S2C0","C1L1"→"S1C1","C1L2"→"S2C1"
(4) : "R1C0"→"C0S1","R2C0"→"C0S2","R1C1"→"C1S1","R2C1"→"C1S2"
(5) : "ES"→"EC0S"
(6) : "R1E"→"E","R2E"→"E"
```

Le premier groupe ne dépend pas de la MT considérée : il traduit le mot d'entrée, w, de la MT dans le langage du SMV sous la forme d'une suite de caractères, c0 (= •) et c1 (= ■). Dans le détail, l'instruction, 1.1, initialise le SMV et les instructions, 1.2 et 1.3, sont appliquées un certain nombre de fois selon la valeur binaire que l'on veut conférer à la donnée initiale. L'instruction 1.4 traduit le fait que la MT est initialisée dans l'état, S1, et que le délimiteur, E, définit le début de la zone interdite à la tête de la MT. Au bilan, la bande de lecture contient deux copies en miroir des données, séparées par les caractères, S1E, Voici le premier théorème qui traduit l'état initial de la MT dans le langage du SMV :

$$E\underbrace{C1C0C1C1}_{w}S1E\underbrace{C1C1C0C1}_{\overline{w}=w_{miroir}}$$

Le deuxième groupe d'instructions traduit mot à mot la suite d'instructions de la MT.

Les troisième et quatrième groupes traduisent les déplacements de la tête de lecture de la MT respectivement vers la gauche et vers la droite.

Le cinquième groupe contient une seule instruction qui a pour fonction de rajouter des caractères, c0 (= •), à l'extrême gauche de la bande au fur et à mesure des besoins afin de lui garantir son infinitude potentielle.

Le sixième groupe traduit la manœuvre d'arrêt de la MT qui s'enclenche dès que le caractère, E, le plus à droite est lu. On note que lorsque l'initialisation de la MT est terminée, il n'y a plus place pour le moindre indéterminisme : la production des théorèmes se poursuit jusqu'à l'arrêt. Au terme de son exécution, la MT présente à gauche du délimiteur, E, le résultat de son exécution et à droite, $\overline{w}$, la copie miroir de la donnée initiale qui a été conservée. Celle-ci n'a, de fait, jamais subi d'altération puisqu'elle se trouve en zone interdite d'accès pour la tête. A tout calcul effectué par la MT correspond la dérivation d'un théorème comme dans l'exemple ci-après. On a souligné, à chaque étape de la transformation, le groupe de caractères qui est justiciable d'une règle de production du SMV :

$$EC1C0C1\underline{C1S1}E\overline{w} \rightarrow EC1C0C1\underline{C1L1}E\overline{w} \rightarrow EC1C0\underline{C1S1}C1E\overline{w} \rightarrow EC1C0\underline{C1L1}C1E\overline{w}$$
$$\rightarrow EC1\underline{C0S1}C1C1E\overline{w} \rightarrow EC1C1\underline{R2C1}C1E\overline{w} \rightarrow EC1C1\underline{C1S2}C1E\overline{w} \rightarrow EC1C1C1\underline{R2C1}E\overline{w}$$
$$\rightarrow EC1C1C1\underline{C1S2}E\overline{w} \rightarrow EC1C1C1C1\underline{R2E}\overline{w} \rightarrow EC1C1C1C1E\overline{w} \equiv EC1C1C1C1EC1C1C0C1$$

Ignorant les intermédiaires de calcul qui ne possèdent aucune valeur de vérité, il ne subsiste que des théorèmes de la forme générale, $Ew'E\overline{w}$, où w' est la réponse de la MT au mot d'entrée, w, $MT(w) = w'$.

De cette manière, il apparaît clairement que tout calcul effectué par une MT qui s'arrête peut être mimé par un SMV correctement construit. En particulier, si la MT est universelle, il en résulte que le SMV correspondant est lui aussi universel au sens de Turing.

Jusqu'ici, il ne se passe rien de très spécial : nous venons simplement de définir un système universel au sens de Turing de plus dans une liste déjà longue. Cependant ce système calculatoire est aussi un système formel et à ce titre on peut se demander où il se situe dans les diagrammes de complétude et de décidabilité, C-D.

Lorsqu'on s'interroge sur les questions de cohérence et de complétude syntaxique d'un système formel, il importe de s'assurer que celui-ci est équipé d'une notion de négation. Autrement dit, il faut que le système formel possède à la fois une notion de preuve et de réfutation. La cohérence est, en effet, l'assurance qu'il est impossible de prouver et de réfuter à la fois une même proposition ou si l'on préfère de démontrer simultanément une proposition et sa négation. De même la complétude syntaxique est l'assurance que toute proposition bien formulée est prouvable ou réfutable. Définissons, comme suit, la négation au sein du SMV,

$$\neg EbEa = E0Ea \quad (b \neq 0) \qquad \Leftrightarrow \qquad \neg E0Ea = EbEa \quad (b \neq 0)$$

Considérons, à présent, le cas général d'un SMV universel au sens de Turing qui est équipé de cette notion de négation. Ce système est certainement cohérent car jamais il ne prouve une proposition et sa négation. Cela résulte du fait que la MTU associée fournit toujours une réponse unique sur sa bande au moment où elle s'arrête, soit 0, soit, b≠0. Par ailleurs, vu l'indécidabilité du problème de l'arrêt, nous avons la certitude qu'il existe des conditions initiales, disons, a, telles que la MTU ne s'arrête pas. Il en résulte qu'il existe des propositions du type, EzEa, qui ne sont ni prouvables ni réfutables au sein du SMV : ce sont des indécidable de ce système qui est donc syntaxiquement incomplet. En fait, un argument diagonal des plus subtil montre que ce SMV est même essentiellement syntaxiquement incomplets. Détaillons cet argument par l'absurde.

Imaginons qu'il existe un SMV, universel au sens de Turing, qui ne soit pas essentiellement syntaxiquement incomplet. Cela signifierait qu'il existerait une extension de ce SMV, notée, SMV', qui serait syntaxiquement complète suite à l'adjonction de règles de production complémentaires. Considérons, à présent, un programme quelconque qui calcule une fonction, f. L'ensemble des programmes étant récursivement énumérable, ce programme porte nécessairement un numéro entier, n, qui peut lui servir de code personnel. Appliquée à la donnée, x, le programme calcule la valeur, b : f(x) = b. Dans la traduction de la MTU qui effectue ce calcul, cela signifie qu'il existe un mot, a, tel que la proposition, EbEa est prouvable dans l'extension, SMV'. Ce mot, a, ne dépendant que de n et de x, on pourrait tout aussi bien le noter, {n,x}, et on pourrait reformuler de façon équivalente que si le $n^{ième}$ programme imprime, b, quand il reçoit la donnée, x, alors la proposition, EbE{n,x} est prouvable dans SMV'. Vient maintenant la contradiction : si SMV' était réellement syntaxiquement complet, on aurait que toute paire, {n,x}, livrerait nécessairement le théorème, EbE{n,x}. Mais alors il existerait une nouvelle fonction, g(x), telle que :

$$\exists g(x) \begin{cases} 1 \;\; si \;\; E0E\{x,x\} \;\; est \;\; prouvable \;\; dans \;\; SMV' \\ \\ 0 \;\; si \;\; EbE\{x,x\} \;\; est \;\; prouvable \;\; dans \;\; SMV' \qquad (b \neq 0) \end{cases}$$

Aucun numéro entier n'est attribuable à cette fonction, g, car g(n) ne coïncide avec aucune fonction programmable. Or cela est impossible puisque toute fonction programmable doit pouvoir programmée sur une MTU. La seule conclusion possible est que l'hypothèse de départ était fausse : le système SMV ne possède aucune extension complète.

Enfin, il nous reste à montrer que les SMV universels au sens de Turing le sont aussi au sens de Gödel, autrement dit, qu'ils sont capables d'émuler n'importe quel système formel. Cela est heureusement plus facile à démontrer car les théorèmes d'un système formel forment un ensemble récursivement énumérable. Il est donc possible d'attribuer un numéro entier, disons, a, à chacun d'eux. On peut alors construire une MT qui prend cet entier en entrée puis qui se livre à une recherche exhaustive des théorèmes de n'importe quel système formel, Σ, imprimant 1 si la proposition de code, a, est un théorème de ce système et imprimant 0 si c'est la négation de cette proposition qui y est un théorème. Le SMV qui émule cette MT répond à la question posée : il émule Σ et comme nous n'avons posé aucune restriction sur Σ, qui pourrait tout aussi bien coïncider avec le SMV lui-même, cette émulation est bien universelle au sens de Gödel.

Systèmes formels capables d'auto émulation.

Un système, Σ, est capable de s'auto émuler quand il est capable de formaliser en son sein les propositions auto référentielles suivantes,

Dans Σ, on formule, p_1 : "Je, p_1, suis prouvable dans Σ".

Dans Σ, on formule, p_2 : "Je, p_2, ne suis pas prouvable dans Σ".

La première proposition n'est pas intéressante : sous peine de contradiction, elle doit être prouvable dans Σ donc vraie, point final. C'est la deuxième proposition qui est intéressante. Elle affirme qu'il n'existe pas de preuve de p_2 dans Σ. Sous réserve que p_2 soit correctement formalisable dans Σ, seuls trois cas sont possibles : p_2 est prouvable, réfutable ou indécidable. Si, p_2, était prouvable dans Σ, on aurait que Σ serait incorrect puisqu'on y démontrerait une proposition fausse. Si p_2 était réfutable dans Σ, la situation serait tout aussi intolérable puisqu'on aurait réussi à réfuter une proposition valide. Il ne reste qu'une seule possibilité : p_2 est un indécidable de Σ. On notera au passage que l'indécidabilité de p_2 a pour conséquence, dans ce cas précis, qu'elle est vraie!

Il n'est pas évident que ce qui apparaît de prime abord comme une boutade puisse faire l'objet d'une formalisation au sein d'un système formel intéressant. Cette construction formelle est pourtant possible et elle repose sur la technique d'arithmétisation des propositions qui est au cœur des travaux de Gödel. Les ressources de l'arithmétique de Robinson (a fortiori de celles de Peano) suffisent pour réaliser cette performance. Il n'est pas connu si cette condition est également nécessaire mais c'est le point de vue généralement partagé par la communauté scientifique.

L'histoire n'est pas terminée pour autant. On pourrait penser qu'il suffirait d'ajouter p_2 à la liste des axiomes de Σ pour que ce système devienne syntaxiquement complet mais ce n'est pas le cas. Tout ce qu'on réussirait à faire, de cette manière, c'est construire un nouveau système formel, Σ', auquel l'argument précédent s'appliquerait à nouveau en sorte que l'incomplétude syntaxique de Σ est bel et bien essentielle. C'est l'essence du premier théorème d'incomplétude de Gödel.

Le premier théorème d'incomplétude de Gödel.

Tout système formel qui atteint le seuil d'universalité au sens de Gödel est essentiellement syntaxiquement incomplet. Personne ne sait où se situe exactement ce seuil mais le fait est que l'arithmétique de Robinson l'atteint alors que celle de Pressburger ne l'atteint pas. Tout ce qu'on pourrait faire, pour en savoir plus à ce sujet, c'est essayer de faire disparaître l'universalité en simplifiant l'axiomatique de Robinson ou de la faire apparaître en complexifiant celle de Pressburger. Vu l'infinité des variantes axiomatiques possibles, il y a peu de chance que ce travail d'encadrements successifs connaisse jamais une fin. Le premier théorème d'incomplétude de Gödel (1931) peut s'énoncer comme suit :

« Il n'existe pas d'extension syntaxiquement complète et cohérente de l'arithmétique de Robinson".

La démonstration rigoureuse que Gödel a donnée du premier théorème est complexe du fait qu'elle veut être constructive en mettant en évidence une proposition indécidable de la théorie. Nous nous contenterons d'une approche intuitive, due pour l'essentiel à Hofstader, qui suffit à suivre la démarche intellectuelle. Le point de départ de la démonstration est la proposition auto-référentielle, p, formulée dans Σ :

p : "Je ne suis pas prouvable dans Σ".

Voici, en gros, comment Gödel a arithmétisé cette proposition. L'idée maîtresse est de transcrire l'énoncé, p, dans la forme équivalente,

p : "Il n'existe pas de couple, {p,D}, tel que D soit une démonstration de p au sein de Σ".

On procède en plusieurs étapes.

1) Codage des propositions.
Il existe une infinité de manières de coder une assertion par un entier, la seule exigence étant que toute assertion bien formée possède un code et un seul. Par contre, on tolère que certains codes ne correspondent à aucune assertion. Gödel utilisait un codage "savant" basé sur la décomposition des entiers selon leurs facteurs premiers mais le codage "naïf" suivant, dû à Hofstader, convient tout aussi bien. On commence par associer un entier à deux chiffres à chaque symbole alphabétique du système formel en évitant le zéro. Optons pour le choix arbitraire suivant :

$0 \leftrightarrow 69$	$S \leftrightarrow 12$	$= \leftrightarrow 11$	$+ \leftrightarrow 17$	$\times \leftrightarrow 71$
$(\leftrightarrow 66$	$) \leftrightarrow 99$	$[\leftrightarrow 51$	$] \leftrightarrow 15$	$a \leftrightarrow 88$
$' \leftrightarrow 18$	$\Rightarrow \leftrightarrow 19$	$\forall \leftrightarrow 44$	$\exists \leftrightarrow 33$	$: \leftrightarrow 55$
$\vee \leftrightarrow 41$	$\wedge \leftrightarrow 14$	$\neg \leftrightarrow 96$	$< \leftrightarrow 24$	$; \leftrightarrow 42$

Le nombre de Gödel (G-code) d'un énoncé s'obtient par traduction littérale des symboles qui le composent, lus dans l'ordre naturel, de gauche à droite. Par exemple, la proposition fermée, $\forall a : \neg Sa = 0$, qui est un axiome de l'arithmétique de Robinson possède le G-code, c = 4488559612881169 soit en résumé :

$\forall a : \neg Sa = 0 \ \leftrightarrow$ c = 4488559612881169 = $S_{44885596128811690}$,

où la notation indicée signifie qu'on répète 4488559612881169 fois le symbole "successeur ", S. Cette proposition est évidemment valide puisque c'est un axiome et qu'on attend d'un axiome qu'il soit vrai dans tout modèle.

2) Définition de la fonction primitive récursive "autosub".
Considérons à présent cet autre exemple, la proposition ouverte, $a = S0$, qui porte sur la variable libre, a. Elle possède elle aussi son code, c :

$a = S0 \ \leftrightarrow$ c = 88111269 = $S_{881112690}$.

Du fait de la présence d'une variable libre, la validité de cette proposition dépend de la valeur particulière de a mais cela n'a pas d'importance pour la suite. Une valeur particulière de a va nous intéresser, précisément celle qui correspond au numéral du code de la proposition qui précède. Cela donne l'énoncé particulier suivant dont le code de Gödel, c', est à son tour évalué :

$S_{881112690} = S0 \ \leftrightarrow$ c' = $12_{881112690}$111269.

L'opération qui consiste à remplacer la variable libre par le numéral du code de la proposition qui l'héberge porte des noms variés dans la littérature. Nous convenons de noter, $c' = autosub[c]$, la fonction qui calcule, c', à partir de la donnée, c.

Cette fonction est primitive récursive puisqu'elle est facilement programmable dans n'importe quel langage primitif récursif : il suffit de traduire pas à pas l'argument informel que nous avons développé sur l'exemple choisi au hasard et d'observer qu'aucune instruction DoWhile n'a été utilisée.

3) Codage des preuves.
Pour développer l'argument de Gödel, nous avons encore besoin d'une autre fonction, également primitive récursive, $m = preuve[n]$, qui calcule le code de Gödel de la première preuve du théorème de code, n, qui se présente lorsqu'on déroule l'ensemble des preuves dans l'ordre canonique.

4) Une étrange proposition.
Considérons à présent la (méta) proposition suivante, contenant la variable libre, a' :

$\neg(\exists a'' : a'' = preuve[autosub[a']])$

En toute rigueur cette proposition sort du domaine d'expression du système formel lorsqu'il utilise les raccourcis commodes que constituent les fonctions, preuve et autosub. La démonstration de Gödel ne s'autorise pas ce genre d'écart car elle désire s'exprimer dans le langage du système. Le prix payé est un développement nettement moins transparent.

Que signifie cette proposition? Qu'il n'existe pas de code de preuve donc de preuve pour l'assertion dont le code résulterait du remplacement de la variable libre présente dans la

proposition de code, a', par cette valeur, a'. Il n'y a rien de particulièrement choquant là-dedans simplement il se peut qu'elle soit vraie pour certaines valeurs de a' et fausse pour d'autres valeurs. En fait, une seule valeur de a' va nous intéresser, celle qui vaut le numéral associé à son code car dans ce cas particulier on aboutit à un résultat étonnant. Cela revient à calculer le code suivant :

$G = autosub[code[\neg(\exists a'' : a'' = preuve[autosub[a']])]]$

C'est le code d'une proposition qui affirme qu'elle n'est pas prouvable dans Σ :

G = "Je ne suis pas prouvable dans Σ" !

Si le système formel dans lequel on travaille est cohérent, il est exclu que G et non-G soient simultanément prouvables. G ne peut pas être prouvable car cela signifierait qu'elle est fausse et néanmoins prouvable ce que personne n'est prêt à accepter. Il ne reste que le cas où G n'est pas prouvable tout en étant vraie! C'est donc un indécidable de ce système formel. L'argument est immédiatement transposable à tout système formel qui permet l'implémentation des fonctions primitives récursives, preuve et autosub.

Une précision est utile à ce stade : le premier théorème d'incomplétude ne se contente pas d'affirmer que l'arithmétique de Robinson ou celle de Peano contient des indécidables, ce serait vraiment trop banal. Il va beaucoup plus loin en affirmant que la théorie est essentiellement syntaxiquement incomplète c'est-à-dire impossible à compléter par ajouts successifs d'axiomes. Après chaque ajout, on se retrouve avec un système un peu moins incomplet certes mais au sujet duquel l'argument conçu par Gödel s'applique à nouveau.

Le programme d'Hilbert de croire en l'existence d'un système formel définitivement syntaxiquement complet est donc une utopie car, au-delà du seuil d'universalité Gödélienne, on a beau ajouter sans cesse de nouveaux axiomes, l'incomplétude syntaxique subsiste.

Le deuxième théorème d'incomplétude de Gödel.

On peut formuler le premier théorème d'incomplétude un peu différemment :

"Tout système formel qui contient l'arithmétique de Robinson est soit essentiellement syntaxiquement incomplet soit incohérent".

Le deuxième théorème de Gödel (1931) précise qu'il n'y a pas moyen de trancher entre ces deux options :

" Il n'y a pas moyen de prouver ni de réfuter la cohérence d'un système essentiellement syntaxiquement incomplet au sein de lui-même, autrement dit, la cohérence d'un système essentiellement syntaxiquement incomplet est un indécidable de ce système "

Par rapport aux propriétés de contradiction et de complétude syntaxique, on a que tout système formel, Σ, qui contient l'arithmétique de Robinson se trouve nécessairement dans l'un des quatre cas suivants :

(1) Σ est syntaxiquement complet et incohérent.
(2) Σ est syntaxiquement complet et cohérent.
(3) Σ est syntaxiquement incomplet et incohérent.
(4) Σ est syntaxiquement incomplet et cohérent.

Le premier théorème d'incomplétude exclut le deuxième cas et il est d'usage d'exclure les cas, 1 et 3, car personne n'est prêt à s'intéresser à un système incohérent où n'importe quoi est prouvable. Il ne subsiste dès lors que la quatrième option : tout système qui contient suffisamment d'arithmétique est syntaxiquement incomplet et cohérent.

Il importe de comprendre qu'on n'a nullement démontré que ces systèmes sont cohérents, on a simplement admis qu'il en était ainsi. Poser que tout système formel qui contient l'arithmétique de Robinson est cohérent s'apparente à un acte de foi auquel on ne peut échapper. Les mathématiques font un usage fréquent du raisonnement par l'absurde. Or ce principe ne fonctionne que si le système dans lequel on raisonne ne peut en aucune manière être tenu pour responsable des contradictions observées! Incorporer à Σ l'axiome, "Σ est cohérent", ne résout rien car cela revient à fabriquer un nouveau système, Σ', qui nous confronte au même problème, déplacé d'un cran.

Gentzen a démontré la cohérence de l'arithmétique de Peano au moyen d'une récurrence transfinie, donc dans le cadre plus puissant de la théorie des ensembles, mais cette démonstration est peu utile car elle postule, à son tour, que la théorie des ensembles est elle-même cohérente. Tout ce qu'on réussit à faire de cette manière, c'est repousser le problème dans un territoire encore moins sécurisé. Le consensus raisonnable est celui qui fait confiance aux axiomes de l'arithmétique élémentaire et de ZF(C) et qui ne pose pas de questions de cohérence à leur sujet.

Equations arithmétiques et diophantiennes.

L'arithmétisation des propositions suffit pour démontrer les théorèmes de Gödel et, en particulier, pour asseoir le pouvoir universel de l'arithmétique de Robinson. A l'époque de leur publication, une majorité de mathématiciens ont considéré que cette technique utilisait un codage tellement alambiqué que, fatalement, les conclusions de Gödel ne devaient pas vraiment les concerner, au moins dans leur pratique quotidienne des mathématiques. Avec le temps, suite aux travaux de Jones, Robinson, Matyiasevic et Chaitin, on a fini par mettre au jour un codage nettement moins exotique appelé diophantinisation. La diophantinisation, qui va nous occuper à présent, est une technique mathématique qui montre en action concrète le pouvoir universel de l'arithmétique élémentaire.

Nous appelons (in)équation arithmétique toute (in)équation ne faisant intervenir que des fonctions primitives récursives d'entiers non négatifs. La limitation aux entiers non négatifs n'en est pas vraiment une puisque toute équation peut être partagée adéquatement entre ses deux membres, comme dans l'exemple suivant,

$1 + 2^m = n^2 + k!$

Cette équation particulière possède des solutions simples, {m,n,k}, faciles à trouver au moyen d'un inventaire progressif et systématique :

```
{{0,0,2},{0,1,0},{0,1,1},{1,1,2},{2,2,0},{2,2,1},{4,4,0},{4,4,1},{5,3,4},{6
,8,0},{6,8,1},{7,3,5},{8,16,0},{8,16,1},{10,32,0},{10,32,1},{11,45,4},…}
```

et il n'est pas interdit de penser qu'elle en possède une infinité. Ce n'est toutefois jamais une mince affaire d'en apporter la preuve car, contrairement à ce qui se passe en algèbre élémentaire des réels, il n'existe aucune procédure de décision valable pour les (systèmes) d' (in)équations arithmétiques.

Il existe une classe particulière d'équations arithmétiques qui ne font intervenir que des polynômes à coefficients entiers : on les appelle (systèmes) d' (in)équations diophantiennes. On pourrait s'étonner d'avoir à s'intéresser à une telle restriction mais cela convient à tous ceux qui veulent faire le lien avec le système formel de Robinson, qui ne construit effectivement que des fonctions polynomiales des variables.

Toute équation arithmétique peut être convertie en une équation diophantienne. A la réflexion, cela n'est pas étonnant puisque l'arithmétique de Robinson est universelle au sens de Gödel. La procédure de conversion s'appelle la diophantinisation de l'équation arithmétique. Elle a toujours pour prix l'introduction de variables supplémentaires et un allongement des écritures. Voici quelques exemples simples de diophantinisations. Pour rappel, les variables, x, y, z, …, a, b, c, …, sont des entiers non négatifs.

On ramène, comme suit, un système d'équations à une équation unique :

$$équation1 = 0 \wedge équation2 = 0 \quad \Leftrightarrow \quad (équation1)^2 + (équation2)^2 = 0$$

$$équation1 = 0 \vee équation2 = 0 \quad \Leftrightarrow \quad équation1 \times équation2 = 0$$

On peut toujours abaisser le degré d'une équation polynomiale en dessous du degré 5 en introduisant suffisamment de variables auxiliaires :

$$x^{13} = y + 1 \quad \Leftrightarrow \quad (a = x^2) \wedge (b = a^2) \wedge (c = b^2) \wedge (d = bc) \wedge (dx = y + 1)$$

Voici comment on diophantinise une inéquation :

$$x < y \quad \Leftrightarrow \quad y = x + c + 1$$

On diophantiniserait, de même, la fonction qui calcule le reste de la division de x par y+1 :

$$z = Mod[x, y + 1] \quad \Leftrightarrow \quad (x = z + (y + 1)a) \wedge (z < y + 1)$$

que l'on transformerait comme suit, au vu des résultats précédents :

$$z = Mod[x, y + 1] \quad \Leftrightarrow$$
$$a^2 + c^2 + x^2 + (a^2 + 1)y^2 + 2z^2 + 2a^2y + 2(a + c)z + 2ayz = 2ax + 2cy + 2axy + 2xz + 2yz$$

Dans cette équation, l'équivalence signifie que ces deux équations arithmétiques possèdent exactement le même ensemble de solutions entières en les variables, x, y et z. On a bien que la deuxième équation est diophantienne. On voit, sur cet exemple pourtant simple, que la diophantinisation s'accompagne d'un allongement substantiel de l'équation de départ. La situation s'aggrave nettement si on diophantinise des équations plus compliquées telles, $z = y^x$ ou $y = x!$, à tel point que nous renonçons à présenter les formes polynomiales exactes correspondantes. La diophantinisation de l'exponentielle requiert à elle seule l'introduction de 57 variables additionnelles !

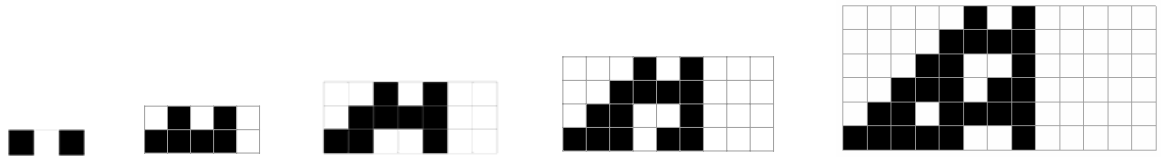
Diophantinisation de la règle 110.

L'automate cellulaire unidimensionnel, 110, peut être arithmétisé puis diophantinisé comme suit. La règle 110 est équivalente à l'opération binaire,

$BitXor[\,BitAnd[\,a,2a,4a\,],BitOr[\,2a,4a\,]]$,

appliquée à l'entier, a, dont la traduction binaire encode le motif de départ, fait de '0' et de '1'. Son application répétée aux nombres qui encodent les états successifs de l'automate fournit l'historique de l'automate 110. Voici, pour l'exemple, cinq pas de l'évolution de l'automate 110 sur la condition initiale, {1,0,1}, que l'on peut assimiler au codage binaire de l'entier, $a=x_2=5$, et qui est immergée dans un environnement vierge, de '0'. A noter que la condition initiale pouvant être plongée dans un environnement vierge ou périodique (voire récursif), il est nécessaire de préciser sa longueur, x_1=width=Length[IntegerDigits[x_2,2]], afin d'éviter toute ambiguïté.

```
{width=3;init={1,0,1};step=5};
ArrayPlot[PadRight[#,width+2 step]&/@CellularAutomaton[110,{init,0},step],Mesh→True]
```



Bien que cela n'ait rien d'indispensable, on s'est arrangé pour que le graphique soit symétrique par rapport à la donnée initiale (c'est le rôle de l'instruction, PadRight). Il en résulte une présence massive de '0' dans les colonnes situées à droite, qui est caractéristique des automates élémentaires qui suivent la règle {1,0,0}→0. Les entiers, x_4, indiqués sous chaque graphique correspondent à la traduction décimale du binaire qui code l'évolution de l'automate. On les a obtenu en lisant, comme dans un livre, la succession des cases noires (='1') et blanches (='0'). Une construction astucieuse mais fastidieuse, due à Wolfram, permet de construire un système arithmétique portant sur 19 variables, x_1=width (=Length[x_2]), x_2=FromDigits[init,2], x_3=step, x_4, x_5, …, x_{19}, qui décrit l'évolution de la règle 110, à savoir :

$$\begin{cases} 1+x_4+x_{12} == 2^{(1+x_3)(x_1+2x_3)}, \;\; 1+x_2+x_{13} == 2^{x_1}, \\ 1+x_5+x_{14} == 2^{x_1}, \;\; 2^{x_3}x_5+2^{x_1+2x_3}x_6+2^{x_1+x_3}x_{15}+x_{16} == x_4, \\ 1+x_{15}+x_{17} == 2^{x_3}, \;\; 1+x_{16}+x_{18} == 2^{x_3}, \\ 2^{(1+x_3)(1+x_1+2x_3)}x_2-x_{10}+x_{11} == 2x_4, \\ (x_7 == BitAnd[x_6, 2x_6]) \wedge (x_8 == BitOr[x_6, 2x_6]), \\ (x_9 == BitAnd[x_6, 2x_7]) \wedge (x_{19} == BitOr[x_6, 2x_7]), \\ (x_{10} == BitAnd[x_9, 2x_8]) \wedge (x_{11} == BitOr[x_9, 2x_8]) \end{cases}$$

Ce système peut posséder un nombre plus ou moins grand de solutions entières. Il est, en fait, construit de telle façon qu'il ne possède qu'une seule solution lorsque l'entier, x_4, correspond très exactement à x_3 pas d'évolution de l'automate 110 sur la donnée, x_2, de largeur, x_1, et il n'en possède aucune si cette condition n'est pas vérifiée. Par exemple, ce système possède la solution unique,

$$\{x_1 \to 3, x_2 \to 5, x_3 \to 1, x_4 \to 350, x_5 \to 7, x_6 \to 10, x_7 \to 0, x_8 \to 30, x_9 \to 0, x_{10} \to 0, x_{11} \to 60, x_{12} \to 673, x_{13} \to 2, x_{14} \to 0, x_{15} \to 1, x_{16} \to 0, x_{17} \to 0, x_{18} \to 1, x_{19} \to 10\}$$

si on impose x_1=3, x_2=5, x_3=1 et x_4=350. Par contre, il n'en possède aucune si on impose, x_1=3, x_2=5, x_3=1 et $x_4 \neq 350$.

Wolfram a diophantinisé le système arithmétique précédent en deux temps. L'élimination des fonctions BitAnd et BitOr mène à une équation diophantienne exponentielle à 79 variables puis l'élimination fait passer ce nombre à 2154 !

Le système arithmétique de Wolfram émule la règle 110 et puisque cette règle est universelle au sens de Turing, il hérite naturellement de cette universalité et par conséquent de son indécidabilité.

Il semble qu'emporté par son élan, Wolfram ait perdu de vue que l'universalité de la règle 110 est de type faible. Ceci a pour conséquence que l'indécidabilité à laquelle elle est sujette n'est sensible qu'asymptotiquement. Le raisonnement qu'il tient selon lequel il est impossible de décider si ce système arithmétique possède ou non une solution lorsqu'on impose les conditions initiales, x_1 et x_2, et un historique, x_4, particulier, ne tient pas la route : il est en effet très facile d'écrire une procédure de décision qui résout ce problème. L'erreur de Wolfram est d'autant plus inexplicable qu'il mentionne lui-même que les automates unidimensionnels ne sont indécidables qu'asymptotiquement.

On peut, en suivant une procédure de codage similaire, traduire n'importe quel automate cellulaire en équation, arithmétique ou diophantienne. On peut, par exemple, diophantiniser le système, Life, et vu que ce système est fortement universel au sens de Turing, l'équation arithmétique correspondante est, elle aussi, universelle.

De même qu'il existe des MTU qui sont capables d'émuler n'importe quelle MT particulière, il existe des équations diophantiennes universelles qui émulent toutes les autres. Ce sont des équations paramétriques de degré fixe qui possèdent les mêmes jeux de solutions que n'importe quelle équation particulière dont le numéro d'ordre est simplement encodé par la valeur entière du (ou des) paramètre, dans une numérotation convenue d'avance.

Ensembles diophantiens et équation diophantienne universelle.

Il peut s'avérer intéressant d'introduire un ou plusieurs paramètres dans l'écriture d'une équation arithmétique. Lorsque c'est le cas, on convient habituellement de réserver les lettres de la fin de l'alphabet aux variables et celles du début aux paramètres. Variables et paramètres sont des entiers non négatifs (= appartiennent à N_0).

Un sous-ensemble de N_0 est diophantien s'il existe une équation diophantienne paramétrique, $p(a;x_1,\cdots,x_m)=0$, telle que les éléments de l'ensemble sont les valeurs du paramètre qui autorisent l'existence d'au moins une solution à l'équation diophantienne. Voici, pour l'exemple, quelques ensembles diophantiens et l'équation paramétrique associée :

- L'ensemble des entiers carrés parfaits, $\{0,1,4,9,\cdots\}$: $x^2=a$

- L'ensemble des entiers non premiers ≥4, $\{4,6,8,9,10,\cdots\}$: $(x+2)(y+2)=a$

Il est généralement moins évident de montrer que les ensembles complémentaires dans N_0 sont aussi diophantiens, à supposer que ce soit le cas :

- L'ensemble des entiers non carrés parfaits, $\{2,3,5,6,7,8,10,\cdots\}$:
$(a-z^2-x-1)^2+(z^2+2z-a-y)^2=0$

- L'ensemble des nombres premiers est lui aussi diophantien mais l'équation associée, découverte par Jones, Sato, Wada et Wiens, requiert un paramètre et 25 variables !

Les sous-ensembles diophantiens de N_0 sont récursivement énumérables : il suffit d'attribuer aux m variables les valeurs, 0, 1, 2, …, en respectant un ordre canonique (par exemple en procédant par sommes de valeurs croissantes), et de ne retenir que les valeurs positives de $a[1-p(a;x_1,\cdots,x_m)^2]$. Un théorème essentiel qui repose sur les résultats d'un travail initial de Robinson, Davis, Putnam, ultérieurement complété par Matijasevic, prouve que, inversement, les ensembles diophantiens sont récursivement énumérables, d'où il résulte que ces deux notions se recouvrent exactement. Ce théorème s'énonce :

"Tout sous-ensemble récursivement énumérable de N_0, noté W, peut être représenté par une équation arithmétique, en fait diophantienne, c'est la contribution de Matijasevic, telle que :

$$a\in W \quad \Leftrightarrow \quad \exists x_1,x_2,\cdots,x_m : P(a,x_1,x_2,\cdots,x_m)=0\ ".$$

La conséquence la plus importante de ce théorème est l'existence d'une équation diophantienne universelle (en fait d'une infinité) qui est à l'équation diophantienne simple ce que la MTU est à la MT simple. Voici comment s'en convaincre.

Il existe, à coup sûr, une MT particulière qui est capable de reconnaître effectivement les éléments de tout sous-ensemble, W, récursivement énumérable (= semi décidable). On peut donc associer une équation diophantienne à chacune de ces MT. Or, les sous-ensembles récursivement énumérables de N_0 sont en infinité dénombrable et ils sont listables sous la forme, W_1, W_2, W_3, …, chacun étant reconnu par une MT particulière. Vu qu'il existe une MTU qui est capable de simuler le comportement de n'importe quelle MT, il en résulte qu'il existe une équation diophantienne qui est capable d'en faire autant vis-à-vis de ses consoeurs, précisément celle qui satisfait la relation,

$$a \in W_n \qquad \Leftrightarrow \qquad \exists x_1, x_2, \cdots, x_m : P(n, a, x_1, x_2, \cdots, x_m) = 0$$

Dans une équation de ce type, n joue le rôle du numéro du programme que l'on veut émuler et les autres paramètres encodent les instances du problème correspondant. Les variables x_i ne sont que des intermédiaires de calcul.

Il n'est pas question de reproduire ici une équation diophantienne universelle in extenso, elle sont beaucoup trop longues et leur intérêt n'est que théorique. Qu'il suffise de savoir qu'on connaît une équation de degré 4 comprenant 58 variables. Si on n'exige pas une forme purement diophantienne, on peut se contenter d'une forme arithmétique due à Jones :

$$(yuz^2 + \xi = (x - ab)q^2) \wedge (q = z^{5^{60}}) \wedge (\lambda + q^4 = 1 + \lambda x^5) \wedge (\theta + 2b = x^5) \wedge (u = c + t\theta) \wedge$$
$$(y = d + v\theta) \wedge (z = 2^w) \wedge (s = q^{16}) \wedge \eta s^2 = (2r)!/(r!)^2) \wedge (r = [z + yq^3 + uq^5 + 2(y - b\lambda) \times$$
$$(1 + ax^5 + z)^4 + \lambda z^5 + \lambda z^5 q^4)q^4][s^2 - s] + [s^2 - 1][q^3 - xu + u + \theta\lambda q^3(z^5 - 2)q^5])$$

Elle contient 14 inconnues et 4 paramètres, a, b, c et d. Elle n'est pas diophantienne eu égard à la présence d'une exponentielle et de deux factorielles. Ce sont les opérations de diophantinisation et de réduction du degré qui exigent les variables et les paramètres supplémentaires. Seules les équations diophantiennes qui correspondent à des MT qui s'arrêtent peuvent être décidées en terme de l'existence d'au moins une solution, les autres ne sont que semi décidables. Il en résulte que le dixième problème de la liste de Hilbert reçoit une réponse négative.

Le dixième problème de Hilbert.

Le dixième problème de Hilbert est le premier exemple concret d'un schéma de problèmes raisonnablement indécidable. Il s'énonce comme suit :

"Trouver un algorithme capable de décider, dans tous les cas, si une équation diophantienne possède ou non (au moins) une solution (entière)".

C'est bien un problème de décision puisque son énoncé ne demande qu'une réponse affirmative ou négative et nullement d'afficher la ou les solutions éventuelles. Toutefois, ceci ne constitue en rien une restriction car de deux choses l'une : soit la réponse est négative et le problème est réglé soit elle est positive et on a l'assurance qu'une recherche exhaustive sur les entiers aboutira en un temps fini.

La réponse à la question posée par Hilbert est négative parce qu'il existe un grand nombre de sous-ensemble de N_0 qui ne sont pas récursivement énumérables : ce sont très précisément les complémentaires des sous-ensembles récursivement énumérables qui ne sont pas récursifs. Le $10^{ème}$ problème de Hilbert est donc indécidable dans N_0. Il est évidemment facile de trouver une solution à une équation diophantienne qui en possède une puisqu'il suffit d'essayer tous les jeux d'entiers dans l'ordre canonique. Mais si la procédure semble s'éterniser, est-ce parce qu'on cherche quelque chose qui n'existe pas ou parce qu'on est trop impatient? On est dans le cas difficile du complémentaire d'un ensemble récursivement énumérable.

Comme toujours, l'indécidabilité ne concerne que l'ensemble de toutes les instances du problème de Hilbert. Il est parfaitement possible de résoudre, au coup par coup, certaines classes d'équations diophantiennes. Tel est certainement le cas des équations d'ordre un et deux ou des équations homogènes pour lesquels une procédure de décision existe.

On ne sait pas grand chose de la frontière entre décidabilité et indécidabilité dans le dixième problème d'Hilbert : on le sait indécidable dans le cas général où aucune restriction n'est posée quant au degré où au nombre d'inconnues mais le seuil exact n'est pas connu. Les systèmes diophantiens à une seule inconnue de degré un ou deux sont décidables et certainement indécidables à partir du degré quatre. Pour trois on ne sait pas. C'est pire en ce qui concerne le nombre d'inconnues : on sait seulement que le problème est décidable dans le cas d'une seule inconnue et indécidable à partir de neuf. Le seuil d'universalité n'est a fortiori pas connu.

Diophantinisation des propositions formelles.

En un sens qu'il va falloir préciser, on peut également proposer un codage diophantien pour les propositions énonçables au sein d'un système formel.

- Le grand théorème de Fermat, $x^n + y^n = z^n$, est une équation exponentielle, que l'on peut transformer en équation diophantienne pure par adjonction de variables auxiliaires. Si le dixième problème de Hilbert avait reçu une réponse positive, on aurait disposé d'une procédure effective capable de décider si cette équation particulière possédait une solution entière. En cas de réponse positive, on aurait même pu trouver une solution particulière au terme d'une recherche exhaustive. Bien entendu, la réponse à la question posée par Hilbert est négative de sorte que la diophantinisation de l'équation de Fermat ne fait pas progresser d'un pas l'étude de ce problème dans le cadre de l'arithmétique.

- La conjecture de Goldbach concernant l'inexistence d'un nombre pair qui ne serait pas la somme de deux nombres premiers peut elle aussi être diophantinisée avec une analyse identique. On procéderait comme suit. Le sous-ensemble des contre exemples éventuels à la conjecture de Goldbach est certainement récursivement énumérable car on peut énumérer les entiers pairs, à partir de 4, et vérifier s'ils sont la somme de deux entiers premiers. Etant récursivement énumérable, ce sous–ensemble est diophantien, d'où l'existence d'une équation diophantienne particulière, $G(a,x_1,x_2,\cdots,x_m)=0$, qui possède une solution, en particulier en terme de a, si la conjecture est fausse et qui n'en possède pas sinon. Autrement dit, la conjecture de Goldbach est équivalente à cette affirmation que, $G(a,x_1,x_2,\cdots,x_m)=0$, ne possède pas de solution entière.

- L'hypothèse de Riemann concernant les zéros de la fonction Zêta peut également être transformée en une équation diophantienne, $R(a,x_1,x_2,\cdots,x_m)=0$, qui ne possède pas de solutions entières si cette hypothèse est vraie et qui en possède sinon. La manière la plus simple d'y parvenir consiste à traduire cette hypothèse en termes de la répartition des nombres premiers, un classique de la théorie des nombres.

En réalité, ces exemples peuvent être généralisés : l'arithmétisation Gödélienne des énoncés formels a précisément pour objet de traduire en langage arithmétique n'importe quelle proposition formelle, même celles ne concernent pas directement un fait arithmétique. Autrement dit, on peut établir le parallélisme suivant :

- de même qu'à toute MTU on peut associer une équation diophantienne paramétrique qui ne possède de solutions entières pour telle valeur, n, de son paramètre que si et seulement si le $n^{ième}$ programme s'arrête,

- à tout système formel universel, on peut associer une équation diophantienne paramétrique qui ne possède de solutions entières pour telle valeur, n, de son paramètre que si la proposition de numéro, n, est prouvable au sein de ce système formel.

Chaitin a résumé comme suit le parallélisme qui existe entre algorithmes et systèmes de preuves :

$$\textit{Données} \quad \Rightarrow \quad \textit{Programme} \quad \Rightarrow \quad \textit{Résultats}$$

$$\textit{Axiomes} \quad \Rightarrow \quad \textit{Preuves} \quad \Rightarrow \quad \textit{Théorèmes}$$

Dans l'ensemble des chaînes de caractères que l'on peut construire sur l'alphabet d'un système formel universel, le sous-ensemble des preuves est décidable comme l'est celui des axiomes. Une preuve est en effet une suite ordonnée d'invocations d'axiomes et de règles d'inférence qui se termine par la citation du théorème prouvé. Le sous-ensemble des théorèmes n'est lui que semi décidable car si on essaye toutes les preuves dans un ordre canonique et que la procédure tarde à faire apparaître le théorème attendu, il est impossible de savoir si cela est dû au fait qu'on cherche quelque chose qui n'existe pas ou si parce qu'on est trop impatient.

Nombres univers et universels.

Au vu de ce qui vient d'être dit, l'indécidabilité du problème de l'arrêt des machines de Turing possède une traduction en termes d'indécidabilité au sein des systèmes formels. Les commentaires suivants s'inspirent des travaux de Chaitin.

Le nombre de Borel. Borel fut le premier à faire la remarque suivante. Considérons l'ensemble, dénombrable, de toutes les chaînes finies de caractères écrites dans un alphabet quelconque, tel l'ensemble, Λ_2, déjà mentionné dans le cas binaire. Ce pourrait tout aussi bien être l'alphabet habituel comportant 26 lettres, complété par le caractère blanc. Parmi toutes ces chaînes, on ne s'intéresse qu'à celles qui posent, en français, un problème de décision sensé, dans quelque domaine que ce soit. Par exemple, dans une logique binaire, la chaîne de caractères «La durée de vie de tout proton libre est infinie » ne peut recevoir que deux réponses, « oui = '1' » ou « non = '0' ». Le nombre de Borel est le nombre ternaire, $0.b_1b_2b_3\cdots$, qui encode les réponses aux questions, prises dans l'ordre lexicographique, en utilisant le chiffre '2' si la chaîne considérée ne correspond à aucune phrase sensée et '0' ou '1' en cas de réponses négatives ou positives. Il est clair que la connaissance de ce nombre apporterait la réponse à toutes les interrogations, pas seulement mathématiques, converties en problème de décision. Ce nombre "univers" n'est évidemment pas calculable. De plus, il est hautement redondant donc compressible car l'information qu'il contient est extrêmement diluée ne serait-ce que suite à un excès dramatique de '2'.

On pourrait similairement définir un nombre de Borel qui ne retiendrait que les questions pertinentes qui se posent dans le cadre d'un système formel particulier.

Le nombre de Turing. Dans le même registre, on peut convenir de ne poser que les questions qui concernent l'arrêt des programmes fournis à une MTU. Par définition, le $n^{ième}$ bit du nombre de Turing, $0.b_1b_2b_3\cdots$, vaut '1' si la MTU s'arrête sur le programme syntaxiquement correct portant le numéro n dans l'ordre lexicographique et il vaut '0' dans le cas contraire. Ce nombre n'est pas calculable parce que s'il l'était, on aurait une solution au problème de l'arrêt, chose que nous savons impossible. On voit sur cet exemple que le point de vue constructiviste s'impose de lui-même : il ne serait d'aucune utilité de prétendre que ce nombre "existe" sous le prétexte que chacun de ses bits ne peut éviter de valoir '0' ou '1'.

Le nombre de Turing est lui aussi largement redondant : la connaissance éventuelle de ses k premiers bits ne constitue en rien k bits d'information car ces bits ne concernent pas des faits indépendants : en réalité ils ne constituent en gros que $\ell g(k)$ bits d'information. L'argument qui justifie cette affirmation est un peu subtil. Un observateur qui souhaiterait obtenir d'un Oracle les k premiers bits du nombre de Turing n'a pas besoin de poser k questions binaires, une par bit à découvrir : il suffit qu'il demande combien de programmes parmi les k premiers s'arrêtent effectivement. La réponse, k', tient évidemment en un entier de longueur $\ell g(k') \leq \ell g(k) << k$ bits. Il suffit alors de faire tourner les k programmes, en parallèle pour gagner du temps, jusqu'à observer l'arrêt de k' d'entre eux. A ce stade on peut interrompre les calculs puisque le compte y est : tous les programmes restants continueront indéfiniment à boucler. Les k bits du nombre de Turing sont donc compressibles d'un facteur, $\ell g(k')/k$. Autrement dit les instances du problème de l'arrêt ne sont pas indépendantes : les nombres de Turing sont bien universels mais la suite de leurs chiffres est largement redondante. Chaitin a découvert qu'il existe d'autres nombres universels qui ne sont quasiment pas redondants : à partir d'un certain rang, toutes leurs décimales déclinent un fait mathématique indépendant.

L'Omega de Chaitin.

Comme le nombre de Turing, le nombre de Chaitin contient l'ensemble des vérités mathématiques mais, cette fois, compressée de façon irréductible. Alors qu'on peut connaître une infinité de chiffres du nombre de Turing, on ne peut connaître qu'un nombre fini de chiffres du nombre de Chaitin qui est fonction de la richesse du contenu axiomatique du système formel étudié. On définit ce nombre comme suit.

Imaginons que nous enfermions dans un grand sac l'ensemble des programmes et que nous en tirions un au hasard. Ω est la probabilité d'arrêt de ce programme. Cette définition n'a de sens que si on précise comment le tirage aléatoire est effectué. La procédure suivante, due à Chaitin, est détaillée par ailleurs, dans l'exposé relatif à la théorie de l'information selon Kolmogorov, et nous n'en rappelons que l'essentiel du principe. Une MTU reçoit un à un des bits tirés à pile ou face. Ils constituent le programme soumis à la MTU qui fonctionne comme un interpréteur, ne consommant les bits successifs que lorsque la tête de lecture les réclame. Ce programme, p^*, devant être auto délimité, il appartient forcément à l'ensemble préfixe,

Λ_{pref}={{10},{110},{111},{01000},{01001},{01010},{01011},{011000},{011001},{011010},{011011},{011100},{011101},{011110},{011111},{0010000000},{0010000001}, …}

Epinglons cette subtilité que les suites appartenant à ce sous-ensemble doivent pouvoir être de longueur arbitrairement grande. Une machine de Turing qui s'arrêterait sur l'ensemble préfixe et complet des programmes '0', '10' et '11' ne serait pas universelle : en effet, aucun

programme plus long ne serait accepté puisqu'il serait obligatoirement préfixé par l'un de ces trois là et que la machine n'explorerait pas plus avant sa bande de lecture. Or on attend d'une MTU qu'elle soit capable de faire tourner des programmes arbitrairement longs.

On arrête les tirages aléatoires dès que leur suite compose un des programmes préfixes, ce qui ne peut manquer de se produire. Nous savons que la probabilité universelle d'occurrence d'une suite préfixe, p*, vaut, $2^{-\ell(p^*)}$, d'où il résulte que Ω s'exprime comme suit :

$$0 < \Omega = \frac{\sum_{p^* \ qui \ s'arrêtent} 2^{-\ell(p^*)}}{\sum_{p^*} 2^{-\ell(p^*)}} = \sum_{p^* \ qui \ s'arrêtent} 2^{-\ell(p^*)} < 1$$

où l'on a tenu compte de l'égalité de Kraft valable pour les codes préfixes complets,

$$\sum_{p^*} 2^{-\ell(p^*)} = 1 .$$

Toute MTU possède son nombre, Ω, et tous ces nombres possèdent les mêmes propriétés dont voici les deux plus importantes.

– Ω est algorithmiquement incompressible donc aléatoire.
La propriété essentielle de Ω est d'être algorithmiquement incompressible. Rappelons que l'on entend par là que les N premiers chiffres de Ω, qui forment l'approximation, Ω_N, ne peuvent pas être calculés par un programme de longueur substantiellement inférieure à N. Voyons pourquoi il doit en être ainsi.

Raisonnons par l'absurde et supposons qu'il existe un programme plus court que N qui soit capable de calculer, avec exactitude, les N premiers chiffres de l'approximation, Ω_N. Il suffit d'ajouter quelques lignes à ce programme pour qu'il effectue en prime le travail suivant : il passe en revue tous les programmes de longueurs inférieures ou égales à un entier quelconque, k, et il les fait tourner pendant k secondes. Au terme de ces exécutions, il prend note de la longueur, ℓ, des programmes qui se sont arrêtés pendant ce laps de temps et il calcule la somme $\sum 2^{-\ell}$ correspondante. Cette somme constitue une première approximation par défaut du nombre, Ω, peu importe qu'elle soit mauvaise. Quand ce travail est effectué, le programme augmente k d'une unité et il recommence le travail complet jusqu'à obtenir une nouvelle approximation par défaut de Ω, sans doute meilleure, en tous les cas pas plus mauvaise. Il réitère la manœuvre autant de fois qu'il le faut jusqu'à tomber, au stade $k=k_N$, sur une approximation dont les N premiers chiffres coïncident avec ceux de Ω_N. Cela ne peut manquer de se produire car la méthode de calcul converge vers Ω et il importe peu si la convergence est effroyablement lente. Arrivé à ce stade, on connaît avec exactitude les programmes de longueurs inférieures ou égales à N qui s'arrêtent. Il est impossible qu'on en ait oublié un car sa contribution ferait qu'on dépasserait la valeur, Ω_N, supposée connue avec exactitude. Avec ce procédé, on aurait réussi à décider le problème de l'arrêt pour tous les programmes de longueurs inférieures ou égales à N à l'aide d'un programme substantiellement plus court, chose que nous savons impossible. En d'autres termes, si un programme est capable de calculer les N premiers chiffres de Ω, il doit nécessairement être de longueur au

moins égale à N. C'est très exactement la définition d'un nombre incompressible ou ce qui revient au même aléatoire. Toutefois Ω est situé à la frontière de l'aléatoire : aucun nombre aléatoire n'est mathématiquement nommable comme il l'est.

Alors que les nombres de Turing sont largement redondants, il faut en gros en connaître 2^n chiffres binaires pour être en mesure de décider quels programmes de longueur inférieure ou égale à n s'arrêtent ou ne s'arrêtent pas, cette redondance disparaît avec les nombres de Chaitin dont les n premiers bits suffisent en gros pour répondre à la même question.

Ω n'est pas calculable.

Si Ω_N n'est, en mettant les choses au mieux, calculable que par un programme de longueur minimale, N, a fortiori, Ω n'est pas calculable dans son intégralité par une procédure effective donc finie. D'ailleurs, si on pouvait connaître toutes les décimales de Ω, on résoudrait du même coup le problème de l'arrêt dans le cas général des programmes de toutes dimensions, ce que nous savons impossible.

Il va de soi qu'il n'est pas question d'espérer calculer les chiffres de Ω en alimentant la bande de lecture de la MTU par tous les programmes de Λ_{pref} dans l'ordre et en ajoutant un terme, 2^{-n}, à Ω, chaque fois qu'un programme de longueur, n, provoque l'arrêt de la MTU. Cela peut fonctionner pour les premières décimales de Ω car on peut prévoir l'arrêt des programmes courts, cela a d'ailleurs été fait par Calude qui a calculé les 64 premiers chiffres binaires du nombre Ω d'une MTU particulière, mais cela échoue dès qu'on considère des programmes de plus en plus longs, précisément à cause de l'indécidabilité de l'arrêt.

On pourrait trouver étrange de prétendre qu'il est impossible de calculer les décimales successives de Ω à l'aide d'un programme de longueur finie alors que la section précédente, relative à l'incompressibilité de Ω, propose une procédure itérative qui semble faire le travail. Il n'en est rien, cependant, car le calcul effectif d'un réel exige que l'on connaisse, à tout moment, le degré de l'approximation atteinte. Or la méthode décrite échoue complètement sur ce point : à aucun moment on n'est sûr des chiffres dont la valeur est acquise définitivement. Cela est dû au fait qu'il existe des tas de programmes qui ne se sont pas arrêtés sans que l'on sache avec certitude s'ils s'arrêteront un jour.

Contenu informationnel des systèmes formels.

Ω est un concentré d'informations mathématiques. Le contenu informationnel du nombre, Ω, est entièrement présent dans la suite de ses chiffres sous la forme générale,

"Le $n^{ième}$ bit de Ω vaut '0' (ou '1')"

Aucune de ces propositions n'est individuellement décodable en terme d'un théorème particulier du système formel, c'est la succession des bits connus qui exploitable. En effet, la connaissance d'un bit supplémentaire de Ω équivaut à la résolution du problème de l'arrêt des MT sur un plus grand nombre de MT et comme toute proposition mathématique peut être

traduite en terme de l'arrêt d'une MT particulière il s'en suit un progrès dans la liste des théorèmes prouvés.

Toutes les propositions sont traduisibles de cette manière, qu'il s'agisse du grand théorème de Fermat, du théorème sur les nombres premiers jumeaux ou de l'hypothèse de Riemann. Connaître les 10000 premières décimales de Ω suffirait, sans doute, à régler le sort de toutes les propositions intéressantes y compris celles qui résistent encore aujourd'hui à toute tentative de démonstration. On se doute que connaître ces décimales est au moins aussi difficile que résoudre tous ces problèmes!

L'incompressibilité de Ω se traduit informellement comme suit dans le cadre de la théorie des systèmes axiomatiques :

"Un kilo d'axiomes indépendants ne peut donner une tonne de théorèmes, tout au plus un kilo".

On exprime cela plus savamment en disant que les théorèmes sont compressés dans les axiomes et qu'on atteint leur complexité algorithmique lorsqu'ils sont maximalement compressés :

$$K(\text{théorèmes}) \approx K(\text{axiomes})$$

Dit autrement, rien n'exclut qu'on calcule les premiers bits de Ω relatifs aux théorèmes d'un système formel : la seule chose qui est impossible c'est d'en calculer davantage que l'ordre de grandeur de la complexité des axiomes avec lesquels on travaille. Les autres bits sont non calculables et c'est la conséquence de la présence de propositions indécidables au sein du système formel. Evidemment, tout se tient, si on ajoute des axiomes, on supprime des indécidables et on gagne la connaissance de quelques bits supplémentaires de Ω.

Dans tout système formel, il y a des propositions, les théorèmes, qui sont vraies en vertu du principe de raison suffisante cher à Leibniz et Hilbert et d'autres qui le sont sans raison! Ces dernières sont des indécidables de la théorie qu'on peut choisir ou refuser d'incorporer dans la liste des axiomes. On en déduit une conception essentiellement dynamique de l'activité mathématique où tout l'art consiste à découvrir de nouveaux axiomes juste assez riches pour créer une extension intéressante du système de départ sans y introduire de contradiction. C'est tout le contraire de la vision statique d'Hilbert, qui croyait inéluctable que l'activité mathématique soit d'essence triviale une fois la base axiomatique posée.

En particulier, il est inutile d'espérer construire un jour un système axiomatique capable de révéler toutes les propositions valides : la théorie de tout n'existe pas en mathématique et encore moins en physique!

Chaitin a montré qu'il était possible de diophantiniser la suite des bits de Ω. Il a construit une équation diophantienne paramétrique sur n qui possède une infinité de solutions entières si le $n^{ième}$ bit de Ω est 1 et qui n'en possède aucune si ce bit vaut 0. La construction de cette équation est fastidieuse, basée essentiellement sur la méthode évoquée d'approximation par défaut de la valeur de Ω. Quant à l'équation diophantienne obtenue par Chaitin, elle ne requiert pas moins de 200 pages!

Incompétence ou indécidabilité?

Il est maintenant possible de dire quelques mots du statut qu'il convient de réserver à celles des conjectures mathématiques qui résistent encore aux efforts des mathématiciens. On ne trouve de telles conjectures qu'au sein des systèmes formels essentiellement syntaxiquement incomplets, ceux qui occupent la zone 1 du diagramme C-D, essentiellement l'arithmétique et ZF(C).

Qu'une démonstration purement arithmétique du théorème de Fermat soit restée hors d'atteinte n'est pas vraiment une surprise quand on voit la longueur de la démonstration de Wiles dans ZF(C). Le système de Peano étant beaucoup plus concis, axiomatiquement parlant, et on peut craindre que la démonstration cherchée soit d'une longueur démesurée. Rien n'interdit cependant, à ce stade, que l'énoncé de Fermat soit un indécidable du système de Peano. On pourrait, dans ces conditions, être tenté d'ajouter l'énoncé de Fermat au système de Peano, ce qui rendrait sa démonstration immédiate. Cette opération serait sans danger car nous savons, grâce à Wiles, que Fermat est vrai et qu'aucun contre exemple n'est à redouter. Cependant, un tel artifice ne devrait duper personne : il équivaut à baisser les bras.

Il arrive pourtant que les mathématiciens aient indirectement recours à ce genre d'expédients. L'exemple de l'hypothèse de Riemann (en fait le huitième problème dans la liste de Hilbert) est célèbre. Sa démonstration n'est pas connue et un grand nombre de théorèmes de la théorie des nombres ne peuvent actuellement se démontrer qu'en l'invoquant, comme dans ces énoncés qui commencent comme suit, "Sous réserve que l'hypothèse de Riemann soit vraie, alors on a …". Tout se passe comme si, faute de mieux, on avait adjoint cette conjecture à la liste des axiomes de ZF(C).

Chaitin a suggéré qu'on ajoute l'hypothèse de Riemann et/ou la conjecture, P≠NP, soit à ZFC soit à Peano mais l'unanimité est loin d'être faite à ce sujet car l'opération demeure potentiellement dangereuse tant que l'on ne connaît pas avec certitude le statut de vérité de ces propositions. Voyons les enjeux sur deux exemples très différents.

L'hypothèse de Goldbach, "Tout nombre pair supérieur à 2 est la somme de deux nombres premiers", n'est pas démontrée à ce jour. Peut-on l'introduire comme axiome dans une extension de Peano? Si elle est prouvée par quelqu'un, disons dans dix ans, alors on n'aura rien fait de mal, tout au plus aura-t-on introduit un axiome redondant dans Peano. Si elle est réfutée, par contre, il faudra l'ôter d'urgence car l'extension construite est contradictoire. Enfin si elle est réellement indécidable, rien de fâcheux n'est à redouter car Goldbach ne pouvant être réfutée, personne ne sera jamais en mesure de lui opposer un contre exemple et l'extension construite sera aussi cohérente que Peano peut l'être.

Ce raisonnement n'est absolument pas transposable au cas de l'hypothèse, P≠NP. Non seulement, personne ne sait actuellement si elle est prouvable, réfutable ou indécidable dans Peano mais personne ne peut déduire de son indécidabilité éventuelle qu'elle serait automatiquement vraie ou fausse. L'incertitude qui plane sur le statut de vérité de l'hypothèse, P≠NP, déconseille de céder à la tentation de l'incorporer aux axiomes de Peano ou de n'importe quel système intéressant.

Modèles non standards.

L'arithmétique de Peano est essentiellement syntaxiquement incomplète. Considérons l'une quelconque de ses propositions indécidables, appelons-la G. Ce pourrait être l'indécidable que Gödel a mis à jour mais n'importe quel autre pourrait faire l'affaire. G étant indépendante des axiomes de base, on peut, au choix, décider de l'ajouter ou d'ajouter sa négation, non-G, aux axiomes existants afin de créer un système légèrement plus puissant que Peano. Dans une logique binaire, G se doit d'être vraie ou fausse pour les entiers standards que nous connaissons. Si elle est vraie (resp. fausse), l'extension qui ajoute G (resp. non-G) aux axiomes de Peano continue de s'appliquer à ces entiers standards : rien ne change en substance, on a simplement construit un système formel un tout petit peu plus puissant que Peano. Par contre, si G est vraie (resp. fausse), l'extension qui ajoute non-G (resp. G) aux axiomes de Peano ne peut plus concerner les seuls entiers standards sous peine d'être en mesure de démontrer des propositions fausses. La manœuvre est cependant a priori autorisée puisque l'axiome ajouté est un indécidable de Peano. La subtilité est que cette extension concerne une classe étendue d'entiers dits non standards.

Si G est vraie et qu'on décide d'ajouter non-G aux axiome de Peano, on crée de toutes pièces une arithmétique non standard (synonyme : non Peanienne). Les entiers de cette nouvelle arithmétique sont les entiers habituels plus des entiers exotiques qui respectent une propriété, à savoir non-G, que nous savons fausse pour les seuls entiers habituels. Bien entendu, les propriétés des entiers naturels n'ont pas changé, seuls des théorèmes exotiques, tels non-G, font leur apparition qui doivent impérativement impliquer des entiers de la nouvelle sorte. Dans la définition de non-G on trouve, en effet, un quantificateur existentiel qui doit maintenant être compris dans un sens élargi, "il existe un entier standard *ou non-standard*, tel que … " et le paradoxe disparaît.

Ces entiers non standards ne sont, à leur manière, pas plus étranges que les "droites" des géométries non euclidiennes.

La théorie des ensembles possède également ses modèles non standards. Le modèle non standard le plus pratiqué est celui de Robinson simplifié par Nelson. Il commence par poser que tout entier est soit standard soit non standard puis il précise les axiomes qui fixent les règles auxquelles ces entiers doivent obéir. Les entiers standards continuent d'obéir aux seuls axiomes habituels de sorte qu'ils coïncident avec les entiers de Monsieur Tout le Monde. Trois axiomes (des indécidables de ZF évidemment) sont ajoutés qui ne concernent que les entiers non standards. Sans entrer dans les détails, commentons informellement l'axiome qui pose que pour tout sous-ensemble fini d'entiers standards, x, il existe un entier, y, tel que : $y>x$. Dans cette extension, l'entier, y, ne peut être standard et on en déduit que les entiers non standards sont plus grands que n'importe quel entier standard, autrement dit infiniment grands. On retrouve, dans un cadre rénové, des considérations déjà évoquées dans le prologue à propos des ordinaux transfinis.

On peut généraliser ce genre d'extension aux réels avec des conclusions similaires. L'inverse d'un réel non standard devient définissable et il apparaît comme infiniment petit. C'est sur ces bases que Robinson a construit une analyse non standard qui assoit rigoureusement le calcul des infiniment petits, si chers aux physiciens.

Epilogue.

Deux conceptions différentes des mathématiques sont défendables mais chacune se heurte, à sa manière, au phénomène d'indécidabilité. Elles sont comme deux versants par lesquels on peut attaquer la montagne. Le versant algorithmique a la préférence des constructivistes et le versant axiomatique a la préférence des formalistes.

Devant un schéma de problèmes, le constructiviste a le réflexe de construire un algorithme dont il espère qu'il viendra à bout de n'importe quelle instance au terme d'une exécution finie. Ce n'est malheureusement pas toujours possible, précisément lorsqu'il y a indécidabilité au sens de Turing, ce que nous avons plutôt appelé absence de calculabilité.

Le formaliste ne s'embarrasse pas de scrupules algorithmiques, il prétend régler en une fois toutes les instances du problème en prouvant un théorème qui court-circuite le calcul effectif de la solution, par exemple un théorème qui affirmerait preuve à l'appui qu'il est inutile de lancer un ordinateur à la recherche d'un nombre pair qui ne serait pas la somme de deux nombres premiers jumeaux. Ce n'est pas non plus toujours possible, précisément lorsqu'il y a indécidabilité au sens de Gödel. La montagne présente bien deux versants d'escalade possibles mais la difficulté de l'un est égale à la difficulté de l'autre et on ne peut échapper à l'indécidabilité de Turing sans tomber sur celle de Gödel.

Au terme de ce périple, le physicien pourrait penser qu'il n'a guère de raison de se vouloir plus catholique que le pape. Si les mathématiciens se sentent très bien dans leur cocon formaliste pourquoi faudrait-il qu'ils se tracassent à changer des habitudes séculaires qui remontent à l'adoption du cadre infinitésimal? Ce cadre, qui s'inscrit dans la logique de la théorie des ensembles est parfaitement défendable. La seule question qui se pose est de savoir s'il n'entraîne pas des dérives qui compromettent une description saine du monde sensible. Lorsque les mathématiciens adoptent l'axiome du choix, ils acceptent sans sourciller qu'on puisse découper une sphère puis recoller les morceaux pour en faire deux à l'identique. Ils sont parfaitement conscients de l'étrangeté de la manœuvre mais elle ne justifie pas à leurs yeux d'abandonner un cadre formel particulièrement riche.

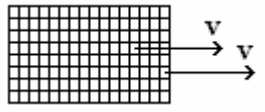
Quand un physicien accepte le cadre infinitésimal, il s'expose également à toutes sortes de dérives telles que des calculs qui mènent à des quantités infinies là où tout esprit sensé attend un résultat fini. Alors qu'une réaction normale consisterait à remettre tout l'ouvrage sur le métier, les physiciens préfèrent s'obstiner dans cette voie en soumettant les résultats des calculs à des bricolages destinés à faire coller ensemble des morceaux qui ne s'adaptent pas. On donne alors un nom savant à la manoeuvre, comme renormalisation, pour faire croire qu'elle est légitime : paresse ou manque d'imagination? De même, dès qu'on écrit des équations différentielles, il faut savoir que certaines possèdent des singularités qui mènent à des comportements étranges. Il ne faut pas chercher bien loin : on sait depuis les travaux de Xia que le mouvement Newtonien tridimensionnel à, $N \geq 5$ corps, peut réserver des surprises pour le moins inquiétantes : il existe des jeux de conditions initiales telles que l'un des corps peut se faire éjecter par les autres et parvenir à l'infini en un temps fini!

On peut certes s'émerveiller qu'une telle chose soit possible mais on peut aussi plus sûrement s'interroger sur le bien fondé d'une théorie qui raconte de telles histoires quand on la pousse dans ses derniers retranchements. Il serait intéressant de voir à quoi la physique ressemblerait si on la resituait dans le cadre constructiviste, en particulier en y remplaçant les équations différentielles par des récurrences.

Nous ne pouvons conclure sans revenir sur cette proposition qu'il est inutile d'espérer déduire d'un ensemble axiomatique de complexité, K, un ensemble de théorèmes de complexité différente de K. Il est en particulier illusoire de prétendre démontrer une proposition qui ne soit pas déjà présente de façon éventuellement obscure dans les axiomes;

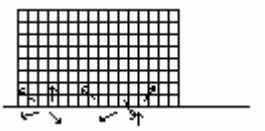
Telle qu'elle, cette affirmation est transposable dans tous les domaines en interdisant de penser qu'on peut créer quelque chose à partir de rien. Ce "quelque chose" ce peut être une organisation et ce "rien" un désordre. Le rapport au second principe de la physique est alors le suivant :

- Lorsqu'un bloc se déplace dans le vide, à vitesse constante, l'encodage des vitesses de chacune de ses cellules élémentaires est vite fait puisqu'elles sont identiques et qu'un seul nombre immuable suffit :



$$K_{fin} = K_{in}$$

- Mais si le même bloc se déplace au contact d'un plan qui le freine jusqu'à l'arrêt complet, les vitesses, celles des cellules du bloc mais aussi celles du plancher, se retrouvent dans un désordre total et leur encodage exige cette fois un programme qui ne peut faire mieux que les citer toutes.



$$K_{fin} > K_{in}$$

Notant, K, la complexité du système à tout instant, soit la longueur du plus court programme qui l'encode, on peut écrire qu'on a toujours, $K_{fin} \geq K_{in}$, et jamais, $K_{fin} < K_{in}$. Autrement dit, il est inutile d'espérer qu'un système isolé gagne spontanément en organisation ou, ce qui revient au même, perde spontanément en complexité.