Le problème des fondations

Nous avons vu que la « révolution moderne » était née du besoin de rigueur. Comme Euclide, qui écrivait son traité, non à l'usage des arpenteurs, mais comme une introduction au monde des « idées » platoniciennes, les créateurs qu'étaient Hilbert, Cantor ou Russell cherchaient à construire un monde purement abstrait dont la cohérence interne était le principal ornement et sa justification même. Les jeunes normaliens qui, dans la décennie précédant la seconde guerre mondiale, mirent en chantier les *Éléments de Mathématique* qu'ils signent en commun du nom mythique de Bourbaki, ne cherchaient pas, sauf accident, à démontrer de nouveaux théorèmes (lorsque l'un des bourbakistes obtient un résultat important, il le publie en général sous son nom personnel avant de l'insérer éventuellement dans le traité) ; leur démarche avait (et a encore, pour leurs successeurs actuels) comme fin principale l'établissement d'un texte de référence qui contiendrait, avec la rigueur maximale, l'énoncé et la démonstration des principaux résultats des mathématiques classiques et modernes [1].

Il n'est pas absolument nécessaire que le mathématicien ait vérifié, par lui-même, la cohérence de toute sa discipline : il peut faire confiance à ses aînés et aux spécialistes pour

1. Il faut se garder d'assimiler mathématiques modernes et bourbakisme. Que Bourbaki soit un des principaux défenseurs des méthodes contemporaines est évident. Il a beaucoup contribué à diffuser l'idée que la plus grande partie des mathématiques doit se ramener à la théorie des ensembles. Mais Bourbaki, même à ses débuts, était beaucoup plus jeune que les mathématiques

Les mathématiciens B. Russell (l'un des pères de la théorie des ensembles), Bourbaki (représenté ici par H. Cartan, J. Dieudonné et L. Schwartz), et A. Lichnerowicz (responsable d'une réforme pédagogique fondamentale).

cela. Ceci est a fortiori vrai du simple utilisateur (comme un ingénieur). Nous voudrions néanmoins donner, en quelques pages, une idée de l'architecture des fondements des mathématiques.

La logique des propositions

Commençons par quelques notions de logique formelle. On peut discuter le fait que la logique appartienne ou non aux mathématiques. Elle précise simplement les règles employées, inconsciemment la plupart du temps, dans la construction d'une science déductive. La logique étudie les relations fondamentales entre propositions prises dans leur plus grande généralité. « P » sera considéré ici comme représentant l'une des expressions ayant un sens mathématique telles que « 5 est un nombre premier », « tous les groupes sont des corps », « il existe un cercle de rayon 1 et d'aire 1 », etc. Nos exemples montrent bien que la logique étudie aussi bien les propositions vraies (« 5 est premier ») que les fausses ; on peut toujours associer à une proposition P l'adjectif vrai (on note alors $v\,(P) = 1$) ou l'adjectif faux ($v\,(P) = 0$), sans que les deux possibilités puissent se présenter ensemble. Peu importe qu'il soit parfois très difficile de calculer effectivement la valeur $v\,(P)$ d'une proposition donnée, l'essentiel est de ne considérer que des propositions pour lesquelles $v\,(P)$ est théoriquement déterminée.

Nous verrons plus loin que cette condition est impossible à vérifier dans tous les cas. Tenons-nous néanmoins provisoirement à cette hypothèse simple. A partir de deux propositions P et Q, on peut définir de nombreuses autres propositions, dont voici les plus importantes :

– $\overline{P}$ (ou encore non P) est la proposition définie par l'égalité $v\,(\overline{P}) + v\,(P) = 1$; elle est vraie si P est fausse et réciproquement (*négation*).

– (P ou Q) est la proposition qui n'est fausse que si les deux propositions P et Q sont fausses ensembles ; on a encore $v(P \text{ ou } Q) = v(P) + v(Q) - v(P)\,v(Q)$. Le « ou » de la

modernes qui datent du début du siècle ; d'autre part son traité, toujours inachevé, ne contiendra jamais qu'une certaine partie des mathématiques, par la volonté même de ses créateurs, partie constamment à réviser par le mouvement même de la recherche. Si précieux qu'il soit, un document historique n'est pas l'histoire elle-même, bien qu'il puisse parfois contribuer à la modifier.

$v(P)$	$v(Q)$	$v(\overline{P})$	$v(P \text{ ou } Q)$	$v(P \Rightarrow Q)$	$v(P \text{ et } Q)$	$v(P \Leftrightarrow Q)$	$v(P \mid Q)$
1	1	0	1	1	1	1	0
1	0	0	1	0	0	0	1
0	1	1	1	1	0	0	1
0	0	1	0	1	0	1	1

Ces « tables de vérité » donnent les valeurs des principales propositions définies en logique, en fonction des valeurs de $v\ (P)$ et de $v\ (Q)$.

P	Q	R	$A = (P \Rightarrow Q)$	$B = (R \text{ ou } P)$	$C = (R \text{ ou } Q)$	$B \Rightarrow C$	$A \Rightarrow (B \Rightarrow C)$
1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1
1	0	1	0	1	1	1	1
1	0	0	0	1	0	0	1
0	1	1	1	1	1	1	1
0	1	0	1	0	1	1	1
0	0	1	1	1	1	1	1
0	0	0	1	0	0	1	1

On peut vérifier, à l'aide de tables de ce genre, des propositions telles que « $(P \Rightarrow Q) \Rightarrow ((R \text{ ou } P) \Rightarrow (R \text{ ou } Q))$ est vraie pour tous P, Q, R ».
(La dernière colonne ne contient que des 1).

disjonction [1] peut encore s'interpréter par la définition suivante : (P ou Q) est vraie si et seulement si l'une des deux propositions *au moins* est vraie.

– $(P \Rightarrow Q)$ est équivalente à la proposition ((non P) ou Q). Cette *implication*, de valeur $v(P \Rightarrow Q) = 1 - v(P) + v(P)\,v(Q)$ diffère nettement de l'implication du langage courant. Elle signifie en effet que l'on ne saurait avoir à la fois P vraie et Q fausse, ce qui est bien conforme au sens courant du mot impliquer, mais elle est également vérifiée si P est fausse, quelle que soit la valeur de Q. (En effet $\overline{P}$ est alors vraie ce qui entraîne que ($\overline{P}$ ou Q) est vraie). En logique, une proposition fausse implique n'importe quelle proposition (vraie ou fausse) : cette distorsion de l'usage courant est néanmoins peu gênante, et nous la conserverons.

1. Le « ou » qui figure dans (P ou Q) est non exclusif, comme le mot latin *vel*, à la différence de *aut* qui signifie « ou » comme dans la phrase « x est pair ou impair », où l'un exclut l'autre.

– (P et Q) équivaut à non ($\overline{P}$ ou $\overline{Q}$) ; sa valeur est $v(P)\,v(Q)$. La *conjonction* n'est vraie que si P et Q sont vraies ensemble.

– ($P \Leftrightarrow Q$) est définie comme étant une abréviation de (($P \Rightarrow Q$) et ($Q \Rightarrow P$)). La valeur de l'*équivalence* est égale à $v(P \Leftrightarrow Q) = 1 - (v(P) - v(Q))^2 = 1 - v(P) - v(Q) + 2v(P)v(Q)$. Elle est vraie si, et seulement si, P et Q ont même valeur.

– ($P \mid Q$) est définie comme étant ($\overline{P}$ ou $\overline{Q}$), et la valeur de l'*incompatibilité*, qui n'est fausse que si P et Q sont vraies ensemble, est $v(P \mid Q) = 1 - v(P)v(Q)$. Elle permet de définir toutes les autres opérations entre propositions, que nous avons définies à partir de la négation et de la disjonction, puisque l'on a :

$$\text{non } P \Leftrightarrow (P \mid P)\,,\ (P \text{ ou } Q) \Leftrightarrow (\overline{P} \mid \overline{Q})$$

Pour certaines propositions, plus générales que les précédentes, $v(P)$ dépend de certaines variables, comme dans la proposition « x est un nombre premier ». Une telle proposition, notée $P(x)$, est vraie pour $x = 2$, fausse pour $x = 4$. Ceci se rencontre notamment dans le cas de propositions *quantifiées*, c'est-à-dire de la forme :
« il *existe* un x tel que P(x) soit *vraie* » notée ($\exists\ x$) $P(x)$ ou « pour *tout* x, P(x) est *vraie* » notée ($\forall x$) $P(x)$. $\exists$ (il existe) et $\forall$ (quel que soit) sont les quantificateurs existentiel et universel.

Axiomes et règles de déduction

Outre les définitions précédentes, les propositions sont reliées par de nombreuses relations. Certaines d'entre elles, choisies parmi celles qui sont vraies quels que soient les éléments et les propositions qui y figurent, sont appelées « axiomes » de la théorie. Citons un système d'axiomes (Hilbert-Ackermann), que l'on doit naturellement compléter par les définitions données ci-dessus de la négation, de l'équivalence, etc. :

- (P ou P) $\Rightarrow P$
- $P \Rightarrow$ (P ou Q)
- (P ou Q) $\Rightarrow$ (Q ou P)
- ($P \Rightarrow Q$) $\Rightarrow$ ((R ou P) $\Rightarrow$ (R ou Q))
- ($\forall\ x\ P(x)$) $\Rightarrow P(y)$
- $P(y) \Rightarrow$ ($\exists\ x\ P(x)$).

Les autres propositions, également universellement valides (et appelées théorèmes : théorèmes et axiomes sont les *thèses*

de notre théorie), peuvent être déduits de ces axiomes par l'application d'un certain nombre de règles :

– règle d'inférence (*modus ponens*) : si P et $(P \Rightarrow Q)$ sont des thèses, alors Q est une thèse ;

– règle de changement de variables : si l'on remplace chaque occurrence de la variable x d'une thèse $P(x)$ par une variable y, alors $P(y)$ est une thèse ; de même si l'on remplace une proposition Q figurant dans une thèse P par une proposition R.

La cohérence d'un système d'axiomes

On pourrait penser qu'un arsenal aussi complexe et subtil permet de donner une valeur à toute proposition bien formulée (c'est-à-dire écrite suivant un certain nombre de règles précises s'appliquant aux propositions ayant un sens mathématique, et excluant des propositions extramathématiques telles que « tout homme est mortel » ou des propositions dénuées de sens comme « un triangle est un nombre premier »). Le problème se pose de savoir par exemple si, à partir d'un certain ensemble S de propositions tenues pour vraies dans une certaine théorie (par exemple S est l'ensemble des axiomes de Peano, ou celui de la structure de groupe), il est possible de démontrer une proposition telle que (A et $\bar{A}$) où A est une proposition bien formulée de la théorie T : si c'est le cas, on dit que T est contradictoire ; sinon T est cohérente (le système S est cohérent).

Il existe des théories contradictoires : par exemple celle des ensembles $(E, \circ)$ où $\circ$ est une opération pour laquelle il existe deux éléments neutres e et f distincts (en effet on doit avoir $e = e \circ f = f$, d'où $e = f$: si A est la proposition $(e \neq f)$, (A et $\bar{A}$) est démontrable dans T). Étant donné un système S d'axiomes, développé par les moyens de la logique que nous avons présentés ci-dessus, il est donc important de se poser la question de sa cohérence. Le théorème de complétude de Gödel affirme que, si un système est cohérent, on peut en construire un modèle (ensemble où les relations, constantes, etc., figurant dans S reçoivent une interprétation précise relativement concrète [1]) dont le cardinal est d'autant plus faible que S contient moins de propositions. Comme tout système pour lequel il existe un modèle est cohérent, on voit que la question « S est-il cohérent ? » peut être résolue,

1. Cf. page 17 : nous avons construit des modèles pour prouver l'indépendance des axiomes de Peano.

dans l'affirmative, par la construction d'un modèle de S. L'exemple de la structure de groupe de l'ensemble $\mathbf{Z}_2 = \{ 0, 1 \}$ montre que si, à l'aide de la logique et des seuls axiomes des groupes on arrivait à démontrer une proposition absurde telle que (A et $\bar{A}$), celle-ci aurait une interprétation dans $\mathbf{Z}_2$ ce qui est contradictoire avec l'existence de ce groupe.

Les théorèmes de Gödel

Gödel, le plus grand logicien vivant, est également célèbre pour trois autres théorèmes dont le second est primordial. Partant d'un système d'axiomes très simples (celui qui contient les propriétés de l'addition et de la multiplication dans $\mathbf{N}$), il a pu démontrer en 1931 que l'on peut formuler une proposition P uniquement à l'aide des symboles introduits dans le système – c'est-à-dire l'addition et la multiplication ordinaires – telle que ni P ni $\bar{P}$ ne puissent être démontrés uniquement à partir des axiomes choisis : P est une *proposition indécidable* de cette théorie. Gödel étant parvenu à formuler convenablement la proposition Q « dans le développement de la théorie, on ne pourra jamais démontrer l'égalité $0 = 1$ », proposition équivalente à l'affirmation de la cohérence de son système d'axiomes, il put également prouver que l'on ne pourrait pas démontrer que Q était vraie dans sa théorie sans employer de notions étrangères à celle-ci. D'une manière plus générale, le grand théorème d'incomplétude implique que *la cohérence d'un système mathématique ne peut être démontrée à l'intérieur de ce système* [1]. Certains théorèmes d'arithmétique, très simples dans leur formulation, demanderont peut-être la considération d'ensembles infinis très complexes pour pouvoir être démontrés. (Peut-être est-ce le cas du théorème de Fermat ?)

En prouvant *l'existence de propositions indécidables dans toute théorie axiomatique*, Gödel montrait qu'il n'existait pas de système d'axiomes assez riche pour que l'on puisse y démontrer la cohérence de tous les sous-systèmes que l'on pouvait en extraire. Nous verrons qu'une confirmation éclatante de ce résultat fut donnée récemment par Cohen en 1963, à propos de l'hypothèse du continu [2].

1. Il faut néanmoins que celui-ci soit assez riche pour contenir l'arithmétique usuelle, qui fournit, comme nous l'avons vu, le point de départ des recherches de Gödel.

2. Voir P.-J. Cohen, *Set Theory and the Continuum Hypothesis*, W.A. Benjamin Inc., New York, 1966, que nous suivons ici pour l'essentiel.

Des problèmes insolubles

Il est très tentant de rechercher des procédures automatisées pour décider, dans la théorie déduite d'un système S d'axiomes, quelles sont les propositions que l'on peut effectivement démontrer dans cette théorie. (Une telle procédure pourrait être confiée, par exemple, à un ordinateur.) La formalisation très poussée de la logique que nous venons d'évoquer ne rend pas absurde un tel projet. Le problème consistant à définir effectivement un tel programme, connu sous le nom de « problème de la décision » est en fait insoluble (Church, 1936). Même dans le cas de l'arithmétique, où il est théoriquement possible de dresser la liste de toutes les questions bien formulées de cette théorie, il n'existe aucune fonction *effectivement calculable* telle que $f(n) = 1$ si la proposition A_n est vraie, et $f(n) = 0$ si A_n est fausse. Un tel résultat est, comme le théorème d'incomplétude de Gödel, un théorème de limitation des systèmes formels.

Il existe de grandes ressemblances entre ces trois résultats négatifs :

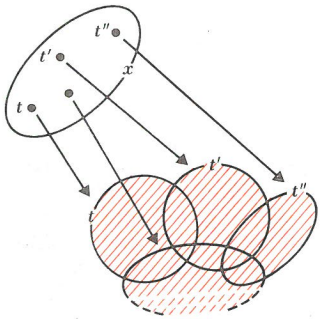
– l'existence de propositions indécidables dans une théorie T, où l'on ne peut démontrer ni P ni $\bar{P}$;

– le fait que l'on ne puisse pas démontrer que le développement de la théorie T ne conduira pas à une absurdité comme (A et $\bar{A}$), du moins tant que l'on reste dans le seul cadre de la théorie ;

– l'existence de problèmes insolubles comme le problème de la décision. Ils peuvent s'interpréter comme des limitations à la compétence d'automates chargés de tirer toutes les conséquences logiques d'un système d'axiomes donné.

Les deux théorèmes de Gödel et le théorème de Church se démontrent d'ailleurs par des méthodes voisines, inspirées de la démonstration par Cantor [1] de l'inexistence d'une bijection entre les ensembles **N** et **R** par la méthode de diagonalisation.

1. Il est fait ici allusion à la démonstration de 1890, exposée par exemple dans *les Nombres et leurs Mystères* (page 125) et non à la démonstration originale plus complexe de 1873.



Le point marqué x est le même objet que l'ensemble $\{a, b, c\}$, la figure de droite est un « agrandissement » de x.

Axiome de la réunion : Si $x = \{t, t', t'', \ldots\}$ l'ensemble hachuré est la réunion des ensembles constituant x : $\cup x = t \cup t' \cup t'' \cup \ldots$
Les flèches constituent un artifice permettant de représenter quand même un objet mathématique à la fois élément (de x) et ensemble dans un diagramme de Venn ; on peut les interpréter comme un « grossissement au microscope », comme dans la figure représentant l'application associant, à x, son successeur x^+.

Les axiomes des ensembles

Dès 1908, le mathématicien Zermelo proposa une axiomatique de la théorie des ensembles, dont nous avons vu le rôle fondamental en mathématiques ; cette axiomatique a été modifiée par Fraenkel. Voici les 9 axiomes, dans une présentation non formelle :

a) *Extension* : deux ensembles sont égaux si et seulement si ils ont mêmes éléments (cf. p. 37).

b) Il existe un ensemble *vide*.

c) Pour tous x et y, il existe la *paire* $\{ x, y \}$.

d) Pour tout ensemble x, il existe un ensemble y, appelé *réunion* de x, tel que pour tout z appartenant à y il existe un élément t de x dont z soit élément et réciproquement (il faut considérer ici x comme un ensemble d'ensembles $t, t', t'' \ldots$ et y comme l'ensemble dont les éléments sont ceux de t, t', t'', etc. ; si $x = \{ t, t' \}$, y est alors $t \cup t'$).

e) *Infini* : il existe un ensemble x tel que, pour tout élément y de x, $y^+ = y \cup \{ y \}$ soit encore élément de x.

f) Pour tout ensemble x, les sous-ensembles y de x sont les éléments d'un nouvel ensemble (l'ensemble $\mathscr{P}(x)$ des *parties* de x).

---

1. Cet axiome bizarre est nécessaire pour éviter des phénomènes tels que l'existence d'un x tel que l'on ait $x \in x$, ou celle de x et de y tels que $x \in y$ et $y \in x$. (Considérer les ensembles $\{ x \}$ et $\{ x, y \}$.) De telles écritures avaient causé de graves soucis aux fondateurs de la théorie des ensembles ; le paradoxe de Russell montrait par exemple que l'« ensemble » des x satisfaisant à $(x \notin x)$ était contradictoire (cf. page 93). L'axiome de régularité écarte évidemment l'existence de l'« ensemble de tous les ensembles ».

g) *Régularité*[1] : pour tout x non vide, il existe un élément y de x tel que $y \cap x = \varnothing$.

h) *Séparation* (ou spécification, sélection, etc.) : étant donné une propriété P et un ensemble x, les éléments y de x qui satisfont à P forment un nouvel ensemble (par exemple $x \cap x'$ est l'ensemble des y de x tels que l'on ait $y \in x'$).

h bis) *Remplacement* : pour certaines questions, il faut substituer cet axiome à l'axiome de séparation qui est trop faible. Il signifie approximativement que si l'on connaît une fonction $u = f(y)$ définie dans l'ensemble x, alors l'ensemble des valeurs u prises par la fonction (l'image de f) forme un ensemble. La notion de fonction utilisée ici est très proche de celle de la page 81 ; une fonction est un ensemble f telle que $(y, u) \in f$ et $(y, v) \in f$ impliquent ensemble $u = v$; simplement il n'est pas précisé a priori si u et v appartiennent à un ensemble.

i) *Axiome du choix*[2] : à tout ensemble non vide x, on peut associer une application f de l'ensemble des sous-ensembles non vides de x dans x lui-même telle que l'on ait $f(y) \in y$.

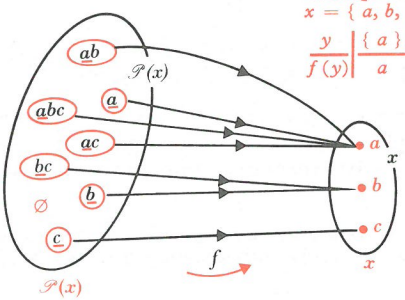
Cet axiome provoqua longtemps de fiévreuses discussions, certains mathématiciens se refusant à considérer des êtres comme l'application f qui n'étaient pas définis de façon explicite, mais il s'est révélé presque indispensable à toute la mathématique.

---

2. Il en existe de nombreuses autres formes ; présenté ainsi, il signifie que l'on peut « choisir » un élément $f(y)$ dans tout sous-ensemble non vide y de x, ce qui n'est pas évident si x est infini. On peut dire de façon légèrement différente (mais équivalente en fait) que pour tout ensemble d'ensembles $\{y, y', y'' \ldots\}$ on peut construire un ensemble z tel que $z \cap y$, $z \cap y'$, $z \cap y''$, etc., ne soient pas vides.

Exemple de fonction de « choix »
$x = \{ a, b, c \}$. f est définie par la table :

y	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{b, c\}$	$\{c, a\}$	$\{a, b, c\}$
$f(y)$	a	b	c	a	b	a	a



On constate que l'on a bien $f(y) \in y$ pour tout $y \subset x$. Dans le diagramme ci-dessus, la lettre « choisie » par f est soulignée.

Ordinaux et nombres naturels

Ces axiomes posés, on peut construire une théorie des *ordinaux* ensembles a totalement ordonnés par la relation $\in$, tels que tout sous-ensemble y non vide admet un élément minimal (c'est-à-dire un élément appartenant à tous les autres éléments de y) et où est vérifiée l'implication

$$(x \in a \text{ et } z \in x) \Rightarrow (z \in a).$$

Les ordinaux sont des ensembles suffisamment nombreux pour que tout ensemble *bien ordonné* (ensemble muni d'une relation d'ordre telle que tout sous-ensemble contienne un élément inférieur ou égal aux autres) puisse être mis en correspondance biunivoque avec un ordinal, la bijection respectant l'ordre défini dans l'ensemble. En d'autres termes, si un ensemble E est bien ordonné par une relation $\leqslant$, il existe un ordinal qui est une copie conforme de E pour tout ce qui concerne la relation d'ordre.

L'importance extrême de cette notion résulte du théorème de Zermelo, équivalent à l'axiome du choix, suivant lequel tout ensemble peut recevoir une relation d'ordre pour laquelle il est bien ordonné [1] ; tout ensemble est donc équipotent à un ordinal au moins. Parmi tous les ordinaux équipotents à E, on en choisit un particulier (celui qui appartient à tous les autres). Cet ordinal particulier est un *cardinal* ; à tout ensemble correspond donc un cardinal unique, qui lui est équipotent. Signalons qu'il n'existe ni ensemble des ordinaux, ni même d'ensemble des cardinaux.

Les nombres naturels sont des cardinaux (et donc des ordinaux) particuliers. Voici comment on peut les définir : 0 n'est autre que l'ensemble vide. A tout naturel x, on associe son successeur $x^+ = x \cup \{ x \}$. Par cette méthode curieuse, on peut ainsi poser :

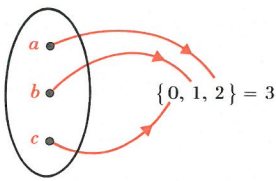
$$0 = \varnothing, 1 = 0^+ = \{ 0 \} = \{ \varnothing \}, 2 = 1^+ = \{ 0, 1 \} = \{ \varnothing, \{ \varnothing \} \},$$
$$3 = 2^+ = \{ 0, 1, 2 \} = \{ \varnothing, \{ \varnothing \}, \{ \varnothing, \{ \varnothing \} \} \}, \text{ etc.}$$

En quelque sorte, les nombres entiers (et par eux une grande partie de la mathématique) reposent uniquement sur... le vide ! Si l'on considère tous les ensembles tels que $\varnothing \in A$ et

1. Si E est bien ordonné, pour définir une fonction de choix f, il suffit de prendre le plus petit élément de chaque sous-ensemble non vide de E (cf. fig. ci-contre).

$$x \in A \Rightarrow x^+ \in A,$$

leur intersection est justement l'ensemble **N** des nombres naturels. On peut y démontrer les axiomes de Peano. Ceux-ci, incontestablement plus simples que la théorie des ordinaux, suffisent à construire l'arithmétique [1].



Tout ensemble fini, bien ordonné (ici $a \leqslant b \leqslant c$) peut être mis en bijection avec un ordinal (ici 3) de telle façon que la correspondance respecte les inégalités dans E en les traduisant dans l'ordinal ($a \leqslant c$ est traduit en $0 \leqslant 2$, etc...)

	$\{a\}$	$\{b\}$	$\{c\}$	$\{ab\}$	$\{ac\}$	$\{bc\}$	$\{abc\}$
min.	a	b	c	a	a	b	a
max.	a	b	c	b	c	c	c

L'ensemble $\{a, b, c\}$ satisfait à la propriété de Stäckel. Si on le munit de la relation $a \leqslant b \leqslant c$, tout sous-ensemble non vide a un maximum et un minimum ; il est donc fini. Par contre, **N** est infini (il existe une injection non bijective de **N** dans **N** définie par $x \rightsquigarrow x^+ = x + 1$).

Un ensemble fini est un ensemble équipotent à un sous-ensemble majoré de **N** (ou encore, avec notre définition de **N**, équipotent à un élément de **N** qui est son cardinal). On peut montrer (Stäckel) qu'un ensemble est fini si, et seulement si, il peut être muni d'une relation d'ordre pour laquelle tout sous-ensemble non vide admet un minimum et un maximum. On peut montrer également, mais en utilisant l'axiome du choix, qu'une condition nécessaire et suffisante pour qu'un ensemble soit infini est qu'il existe une injection de cet ensemble dans lui-même qui ne soit pas bijective (cette définition est due à Dedekind en 1887).

1. On peut aussi utiliser les axiomes suivants : **N** est l'un quelconque des ensembles (tous isomorphes) E, bien ordonnés par une relation $\leqslant$, sans élément maximum mais où tout sous-ensemble majoré non vide admet un maximum.

Soit deux propositions $P(x)$ et $Q(x)$ dépendant d'une variable x décrivant un ensemble E. Notons A l'ensemble des x pour lesquels P est vraie, et B l'ensemble des x pour lesquels Q est vraie.

- A la proposition $\overline{P}$ correspond l'ensemble $\bar{A} = E - A$;
- A la proposition (P ou Q) correspond l'ensemble $(A \cup B)$;
- A la proposition (P et Q) correspond l'ensemble $(A \cap B)$. L'ensemble des propositions $P(x)$ définies sur E est donc une algèbre de Boole isomorphe à $(\mathscr{P}(E), \cap, \cup, -)$. Des théorèmes tels que

$$\overline{\bar{A}} = A,\ \overline{A \cup B} = \bar{A} \cap \bar{B},\ \overline{A \cap B} = \bar{A} \cup \bar{B}$$

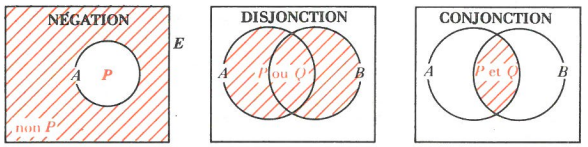
se traduisent en calcul propositionnel par :

$$\overline{\overline{P}} = P,\ \overline{P \text{ ou } Q} = \overline{P} \text{ et } \overline{Q},\ \overline{P \text{ et } Q} = \overline{P} \text{ ou } \overline{Q}$$

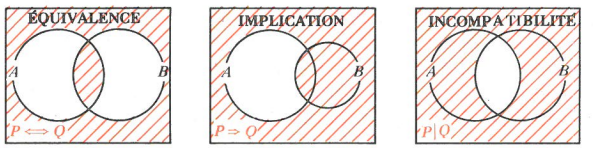
(lois de Morgan).

La fonction v ($P(x)$) est simplement la fonction caractéristique de l'ensemble A, d'où la similitude des formules :

$$\begin{cases} f_{A \cup B} = f_A + f_B - f_A f_B \\ v(P \text{ ou } Q) = v(P) + v(Q) - v(P)\, v(Q), \text{ etc...} \end{cases}$$



- A la proposition ($P \Leftrightarrow Q$) correspond $(A \cap B) \cup (\bar{A} \cap \bar{B})$;
- A la proposition ($P \Rightarrow Q$) correspond l'ensemble $(\bar{A} \cup B)$, ensemble des x pour lesquels P est fausse ou pour lesquels P et Q sont vraies ensemble ; $(\overline{\bar{A} \cup B} = A - B)$.
- A la proposition ($P|Q$) correspond $\overline{\bar{A} \cup \bar{B}} = \overline{A \cap B} = A \perp B$.



- Si P est toujours vraie, cela signifie $(A = E)$. Si $\overline{P}$ est toujours vraie, alors $(A = \varnothing)$. Si $(\exists\ x\ P(x))$ est vraie, alors $(A \neq \varnothing)$. Si $(\forall\ x\ P(x))$ est vraie, alors $(A = E)$.

Si (P ou Q) est toujours vraie, alors ($A \cup B = E$) · Si (P et Q) est toujours vraie, alors ($A \cap B = E$) (c'est-à-dire $A = E$ et $B = E$). Si ($P \Leftrightarrow Q$) est toujours vraie, alors ($A = B$). Si ($P \Rightarrow Q$) est toujours vraie, alors ($A - B = \varnothing$) ou ($A \subset B$). Si ($P | Q$) est toujours vraie, alors ($A \cap B = \varnothing$) (A et B sont disjoints).
Les réciproques de ces propriétés sont vraies.

• On peut interpréter, dans $\mathscr{P}(E)$, les axiomes suivants de Hilbert-Ackermann :
(P ou P) $\Rightarrow P$: $(A \cup A) \subset A$
$P \Rightarrow$ (P ou Q) : $A \subset (A \cup B)$
(P ou Q) $\Rightarrow$ (Q ou P) : $(A \cup B) \subset (B \cup A)$.
et de même
$(P \Rightarrow Q) \Rightarrow ((R \text{ ou } P) \Rightarrow (R \text{ ou } Q))$
se traduit par
$(\bar{A} \cup B) \subset (\bar{C} \cup \bar{A} \cup (C \cup B))$
(si l'on veut éviter l'inclusion, il suffit de remplacer $A \subset B$ par l'égalité $\bar{A} \cup B = E$).

• Le modus ponens équivaut à la remarque suivante : si A et B, sous-ensembles de E, sont tels que l'on ait $A = E$ et $A \subset B$, alors on peut écrire aussi $B = E$.

Des définitions ensemblistes de l'addition et de la multiplication dans N

Nous poserons a priori les axiomes

$$x + 0 = x \qquad x + y^+ = (x + y)^+$$
$$x \times 0 = 0 \qquad x \times y^+ = (x \times y) + x$$

Les deux premiers permettent de montrer, par récurrence, l'existence de $(x + y)$ pour tout x et tout y, ainsi que l'unicité de ce nombre qui résulte de sa calculabilité. On en déduit successivement (par récurrence en général) les égalités ou implications :
$x + (y + z) = (x + y) + z$, $x + 0 = 0 + x$,
$x + 0^+ = 0^+ + x$, $x + y = y + x$,
$(x + y = z + y) \Rightarrow (x = z)$, $(x + y = 0) \Rightarrow (x = y = 0)$,
$x(y + z) = xy + xz$, $x(yz) = (xy)z$,
$x^+ y = xy + y$, $0x = 0$, $xy = yx$, $(x + y)z = xz + yz$.
Définissant ensuite la relation
$$x \geqslant y \Leftrightarrow \exists z, x = y + z$$
on montre que c'est une relation d'ordre total, et que
$(x \geqslant y) \Leftrightarrow (x + z \geqslant y + z)$, $(x \geqslant y) \Rightarrow (xz \geqslant yz)$,
$(xy = 0) \Leftrightarrow (x = 0 \text{ ou } y = 0)$, $(xy = zy) \Rightarrow (y = 0 \text{ ou } x = z)$
$(xy \geqslant zy) \Rightarrow (y = 0 \text{ ou } x \geqslant z)$, $(x \geqslant y) \Leftrightarrow (x^+ > y)$,
$(x \geqslant y \text{ et } y \neq 0) \Leftrightarrow (x > y^-)$, $(xy = 1) \Leftrightarrow (x = y = 1)$;
on peut ensuite définir la division à une unité près, etc.

Les hypothèses du continu

Les ordinaux et les cardinaux infinis sont appelés *nombres transfinis*, et possèdent une arithmétique très particulière. C'est à propos des cardinaux transfinis que se pose le problème du continu. Le théorème de Cantor (voir page 93) montre qu'il existe une injection d'un ensemble E dans son ensemble des parties 2^E, mais aucune injection en sens inverse. Une telle relation définit une inégalité stricte dans les cardinaux, notée card E < card 2^E, inégalité qui joue le rôle de relation d'ordre total dans tout ensemble de cardinaux. En particulier 2^N, qui est équipotent au corps $\mathbf{R}$ des nombres réels, a un cardinal (le « continu ») strictement supérieur à celui de $\mathbf{N}$ (noté aleph zéro et appelé « le dénombrable » : un ensemble dénombrable est un ensemble équipotent à $\mathbf{N}$, comme $\mathbf{Z}$ ou $\mathbf{Q}$). L'*hypothèse du continu*, énoncée en 1878 par Cantor lui-même, pose en principe qu'il n'existe pas d'ensemble E « entre » $\mathbf{N}$ et $\mathbf{R}$, c'est-à-dire encore que tout sous-ensemble infini de $\mathbf{R}$ est équipotent à $\mathbf{N}$ ou à $\mathbf{R}$. L'hypothèse généralisée du continu affirme de même qu'il n'y a pas d'ensemble « entre » E et $\mathscr{P}(E)$. Dans ses travaux sur la recherche d'une démonstration de la non-contradiction de l'arithmétique (ou de la théorie des ensembles), Gödel démontra en 1938 que l'existence d'une contradiction dans la théorie des ensembles, enrichie de l'hypothèse généralisée du continu, entraînerait l'existence d'une contradiction dans les conséquences des axiomes *a* à *h* bis inclus (théorie des ensembles sans axiome du choix). Ceci laissait donc supposer que ces deux propositions, axiome du choix et hypothèse généralisée du continu (et a fortiori l'hypothèse simple elle-même) étaient très vraisemblables et que l'on devait les accepter si l'on croyait aux ensembles, dans la mesure où leur introduction ne pouvait être cause de catastrophes. La situation fut renversée en 1963 par un jeune américain, Paul J. Cohen [1]. Son prédécesseur avait démontré, notamment, que l'axiome du choix et l'hypothèse généralisée du continu étaient des théorèmes pour une classe très vaste d'ensembles, les ensembles « constructibles » : Cohen chercha par exemple un modèle dans lequel il n'était pas vrai que tout ensemble soit constructible. A l'aide d'une méthode originale, il put démontrer à l'aide de différents modèles :

1. Celui-ci reçut, en 1966, la médaille Fields (équivalent du Prix Nobel en mathématiques) pour sa découverte.

– que les huit premiers axiomes de Zermelo-Fraenkel n'impliquent pas l'axiome du choix : il existe des théories des ensembles dans lesquelles celui-ci est faux ;

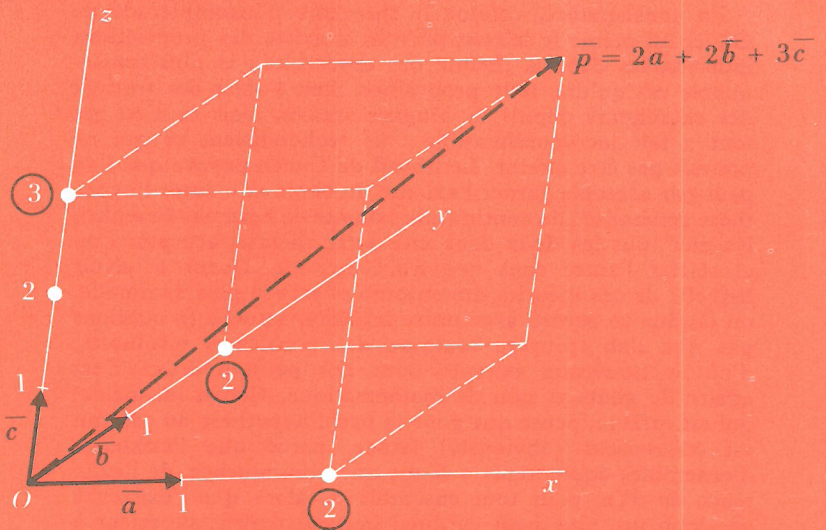
– que les neuf axiomes n'impliquent pas l'hypothèse simple du continu : il existe des théories des ensembles où l'axiome du choix est valable mais non l'hypothèse ;

– que les neuf axiomes et l'hypothèse généralisée n'impliquent pas que tout ensemble est constructible (au sens de Gödel).

Une conclusion très provisoire

En conséquence, comme le théorème d'incomplétude de Gödel lui-même le laissait prévoir, il existe des propositions indécidables dans la théorie des ensembles ; le plus remarquable est qu'une telle proposition était l'objet des travaux des chercheurs depuis de longues années, mais ce n'est pas tout à fait incompréhensible : ils recherchaient ce qui ne pouvait pas être atteint. Le travail de Cohen prouve que l'on peut par exemple garder l'axiome du choix et ajouter l'hypothèse généralisée du continu aux axiomes de Zermelo-Fraenkel, les nier tous les deux dans une autre théorie, accepter l'un et rejeter l'autre dans une troisième, etc. Quant à savoir laquelle de ces théories des ensembles sera la plus commode, ou la plus en accord avec notre intuition actuelle (n'oublions pas que l'on trouvera toujours des problèmes insolubles, d'où de nouveaux axiomes pour nos petits-neveux !) c'est affaire de goûts et non de mathématique. Cohen lui-même, qui est orfèvre, pense que pour sa part l'hypothèse du continu est *évidemment* (obviously) fausse, parce que l'existence d'ensembles équipotents à $\mathbf{R}$ est une conséquence de l'existence de $\mathscr{P}(x)$ pour tout ensemble x, alors que l'existence de aleph un, plus petit cardinal immédiatement supérieur à aleph zéro, résulte d'un procédé beaucoup moins puissant. En d'autres termes, l'axiome h bis (remplacement) est beaucoup plus faible, en ce qui concerne la création de nouveaux ensembles, que l'axiome f qui assure l'existence de $\mathscr{P}(x)$.

Quant à l'axiome du choix, son utilité dans les mathématiques actuelles est si grande (notamment par le biais du théorème de Zermelo et d'un théorème dû à Zorn, qui lui sont équivalents), que l'on ne pense généralement pas pouvoir s'en séparer ; l'avenir dira dans quelles voies fécondes nos successeurs s'engageront sur ces points.



Espace à trois dimensions.
Le vecteur $\overline{p} = 2\,\overline{a} + 2\,\overline{b} + 3\,\overline{c}$ est construit ici comme la diagonale d'un parallélépipède dont les côtés, parallèles aux vecteurs de la base $\{\,\overline{a}, \overline{b}, \overline{c}\,\}$ ont respectivement pour mesures : $2\,\overline{a}$, $2\,\overline{b}$, $3\,\overline{c}$. (2, 2, 3) sont les coordonnées (x, y, z) de ce vecteur $\overline{p}$ (et celles de son extrémité sur cette figure).

Un peu de calcul linéaire

Nous avons déjà rencontré la structure d'espace vectoriel dans l'une des différentes pages du catalogue « algèbre ». Après avoir souligné son importance, nous n'y sommes plus revenus ; peut-être le lecteur qui aurait découvert cette structure dans ce livre (bientôt elle sera commune à tous les lycéens) en a-t-il un vague souvenir comme d'une généralisation mélangeant un groupe et un corps, possédant un modèle économique : le panier de la ménagère.

Une définition très vague, mais qui contient l'idée essentielle, d'un espace vectoriel est celle d'un ensemble où l'on peut additionner des êtres hybrides, nés de la multiplication par un nombre (réel par exemple) d'un « vecteur », c'est-à-dire d'un élément de l'espace vectoriel [1].

En fait, partout où l'on considère des *combinaisons linéaires*, c'est-à-dire des expressions de la forme

$$a\,\alpha + b\,\beta + c\,\gamma + \ldots,$$

constituées de certaines sommes de produits de deux éléments, on travaille presque toujours dans un vectoriel. Le calcul dans un corps commutatif K, qui est un espace vectoriel sur K lui-même, en est le plus simple exemple.

---

1. Le langage géométrique ici utilisé (espace au lieu d'ensemble, vecteur au lieu d'élément) s'explique pour des raisons historiques, voire pédagogiques (au bon vieux temps ou les vecteurs libres étaient « concrets » et les vectoriels abstraits ! D'ici peu, ce sera l'inverse.) Ce langage constitue un obstacle dans la mesure où notre concept est bien plus général.

Les applications linéaires

La notion fondamentale, en algèbre linéaire, est celle d'*application linéaire*. Une telle application, définie sur un vectoriel E, respecte les deux opérations fondamentales : l'addition des vecteurs $(+)$ et la multiplication par un nombre. En d'autres termes, on doit avoir pour tous vecteurs $\vec{p}$, $\vec{q}$ et tout nombre x, les égalités :

$$\begin{cases} f(\vec{p} + \vec{q}) = f(\vec{p}) + f(\vec{q}) \\ f(x\,\vec{p}) = x\,f(\vec{p}). \end{cases}$$

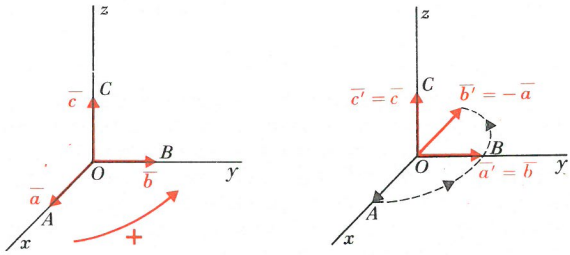
Elles impliquent naturellement que l'application transforme un vecteur de E en un élément $f(\vec{p})$ d'un espace vectoriel, que nous noterons F, puisque l'on doit pouvoir y effectuer les mêmes opérations que dans E : ajouter, et multiplier par un nombre. L'application qui associe, à une fonction dérivable, sa dérivée est une application linéaire de l'espace des fonctions numériques dérivables E dans l'espace plus vaste des fonctions numériques F.

Cet exemple mettait en jeu deux espaces de dimension infinie : le cas le plus simple est celui où E est de dimension finie ; on peut toujours supposer alors que F est lui-même de dimension finie, car l'image d'un espace vectoriel de dimension n par une application linéaire est elle-même un espace vectoriel de dimension au plus égale à n. Supposons en effet $n = 3$, comme dans l'espace géométrique classique ; cela signifie que l'on peut trouver une base $\{ \vec{a}, \vec{b}, \vec{c} \}$ de trois vecteurs qui engendrent tous les autres. Chaque vecteur p peut s'écrire, d'une façon et d'une seule, sous la forme

$$\vec{p} = x\,\vec{a} + y\vec{b} + z\vec{c}.$$

Son transformé $f(\vec{p})$ est alors égal à $\vec{p}' = x\vec{a}' + y\vec{b}' + z\vec{c}'$, où $\vec{a}' = f(\vec{a})$, $\vec{b}' = f(\vec{b})$, $\vec{c}' = f(\vec{c})$ sont trois vecteurs de F. Ou bien $(\vec{a}', \vec{b}', \vec{c}')$ est un triplet de vecteurs indépendants, ce qui signifie qu'aucun d'eux ne peut s'écrire, par exemple, sous la forme $\vec{c}' = u\,\vec{a}' + v\,\vec{b}'$; on montre alors facilement que tout vecteur de $f(E)$ peut s'écrire de façon unique comme combinaison des vecteurs $(\vec{a}', \vec{b}', \vec{c}')$ qui forment alors une base de l'espace $f(E)$ qui est bien vectoriel (et de dimension 3). De plus $\vec{p}$ et $f(\vec{p})$ ont mêmes *coordonnées* (x, y, z) dans les deux bases ; f est donc un isomorphisme de E sur $f(E)$ pour la structure d'espace vectoriel, c'est-à-dire que l'on peut traduire toute relation dans E par une relation analogue dans l'image $f(E)$, et réciproquement.

Si $\vec{a}'$, $\vec{b}'$ et $\vec{c}'$ ne sont pas indépendants, on montre que



Exemple d'isomorphisme linéaire. Rotation d'axe Oz. ici $E = F$ est l'espace euclidien, rapporté à 3 droites Ox, Oy, Oz, deux à deux perpendiculaires, portant des vecteurs de longueur 1 : la rotation f d'axe Oz et d'angle 90° ($\pi/2$) laisse invariant $\bar{c}$ et Oz, et transforme $\bar{a}$ en $\bar{a}' = \bar{b}$, $\bar{b}$ en $\bar{b}' = -\bar{a}$.

l'on peut en extraire deux, un ou même zéro vecteur qui constituent une base de $f(E)$, qui est alors un espace vectoriel de dimension 2, 1 ou même 0 ($f(E)$ est alors formé d'un seul vecteur, le vecteur nul $\bar{o}'$ de F).

Morphisme ou isomorphisme ?

Précisons ces notions par un exemple. E sera ici l'ensemble des polynômes de degré 2 au plus, $\bar{p} = x\,\bar{a} + y\,\bar{b} + z\,\bar{c}$, dont les coefficients (x, y, z) appartiennent au corps $\mathbf{R}$; la base la plus simple de cet ensemble de polynômes, dont nous noterons X la variable, est naturellement la base :

$$\bar{a} = X^0 = 1,\ \bar{b} = X,\ \bar{c} = X^2$$

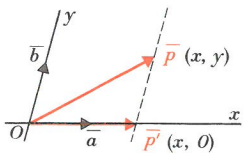
d'où la forme traditionnelle $\bar{p} = x + y\,X + z\,X^2$. A ce polynôme nous associons le vecteur $\overrightarrow{OM}$ de l'espace euclidien F dont les coordonnées sont (x, y, z). Cette application est bien linéaire, et c'est même un isomorphisme de E sur F, puisque ici $f(E) = F$; $\bar{a}' = f(\bar{a})$ est le vecteur $\overrightarrow{OA}$ dont les coordonnées sont (1, 0, 0), etc.

Si nous considérons au contraire l'application f qui, au polynôme $(x + y\,X + z\,X^2)$ associe le polynôme dérivé [1] $(y + 2z\,X)$, celui-ci appartient bien à un espace vectoriel F ; on peut même prendre $F = E$ en écrivant $y + 2z\,X = y\,\bar{a} + (2z)\bar{b} + 0\,\bar{c}$. Toutefois les vecteurs $\bar{a}' = 0$, $\bar{b}' = \bar{a}$ et $\bar{c}' = 2\bar{b}$ ne sont plus indépendants, puisque $\bar{a}' = 0\bar{b}' + 0\bar{c}'$. L'application linéaire que nous avons définie n'est pas un isomorphisme de E sur lui-même ; elle transforme E en

1. Le polynôme dérivé d'un polynôme P est obtenu en remplaçant systématiquement X^0 par 0, et X^n par nX^{n-1} $(n \geqslant 1)$

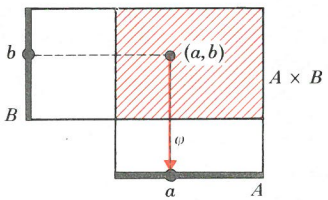
$f(E)$, espace de dimension 2 des polynômes de degré au plus égal à 1, strictement inclus dans $F = E$. Sans être bijective, f respecte néanmoins en partie la structure d'espace vectoriel de E, puisqu'elle transforme une somme en somme et un produit en le produit correspondant ; on dit que c'est un *morphisme*, c'est-à-dire une traduction fidèle dans un sens, mais non nécessairement réversible.

La différence essentielle entre les isomorphismes d'espaces vectoriels et les simples morphismes est que, dans le premier cas, l'équation $f(\overline{p}) = \overline{0}'$, où $\overline{p}$ est un vecteur inconnu de E et où $\overline{0}'$ est l'élément neutre de l'addition dans F, n'a qu'une solution et une seule : $\overline{p} = \overline{0}$ (élément neutre de E). Le morphisme le plus général est moins simple ; les vecteurs $\overline{p}$ de E dont le transformé est nul forment tout un espace vectoriel dans E (on dit un sous-espace vectoriel de E, étant bien entendu que les opérations et le corps K sont les mêmes).



Exemples d'application non bijective. Dans un plan, on projette le vecteur $\overline{p}$ sur l'axe portant $\overline{a}$: $\overline{p}' = f(\overline{p})$ est obtenu en menant par l'extrémité de $\overline{p}$ la parallèle à $\overline{b}$. Le noyau est la droite Oy, l'espace-image $f(E)$ est la droite Ox.

Cet espace vectoriel est le *noyau* de l'application f. C'est l'ensemble des vecteurs qui sont « dissous », ou réduits à zéro, par f. L'exemple le plus simple de morphisme non bijectif est celui d'une projection. A tout couple (x, y) associons le couple $(x, 0)$. Si x et y sont des éléments d'un même corps K, le produit $K \times K$ est l'espace vectoriel le plus simple de dimension deux sur K : on le note évidemment K^2. La projection $(x, y) \rightharpoonup (x, 0)$ le transforme en un sous-espace vectoriel de dimension 1. Le noyau est l'ensemble des couples $(0, y)$ qui sont bien transformés en $(0, 0)$ qui joue ici le rôle de 0 de K^2. L'image géométrique sous-jacente est bien connue (au lieu de vecteur $\overrightarrow{OM}$, on y parle néanmmoins plus souvent du point M lui-même, ce qui n'est guère gênant ici) : au vecteur de coordonnées (x, y), on associe le vecteur $(x, 0)$ obtenu par projection sur $\overrightarrow{Ox}$ parallèlement à $\overrightarrow{Oy}$; le noyau est l'ensemble des vecteurs parallèles à $\overrightarrow{Oy}$, $\overrightarrow{Ox}$ étant l'espace-image.



Une projection. Le concept général de projection est le suivant : à tout élément du produit cartésien $A \times B \times C \times \ldots$ où $A, B, C, \ldots$ sont des ensembles donnés, on fait successivement correspondre les éléments $a, b, c \ldots$ formant le n-uple $(a, b, c, \ldots) \in A \times B \times C \times \ldots$ Ces applications sont les projections du produit sur les ensembles A, B, C, etc.

L'application φ définie par $\varphi(a, b) = a$ est la projection de $A \times B$ sur A. Si B contient un élément o, on appelle encore projection de $A \times B$ sur A l'application de $A \times B$ sur $A \times \{o\}$ définie par $\varphi'(a, b) = (a, o) \in A \times \{o\}$.

Un peu d'espionnage

Notre objet n'est évidemment pas de faire un cours d'algèbre linéaire, les ouvrages scolaires étant là pour cela. Mais ces quelques définitions (dans un livre qui n'en manque déjà pas !) étaient nécessaires pour nous amener au seuil d'un chapitre très particulier, mais important, de l'algèbre linéaire, le calcul matriciel. Comme tout calcul, ce n'est qu'un moyen et non un but pour les mathématiciens qui tentent toujours d'y substituer « des idées ». Mais un calcul a aussi l'avantage d'être plus concret, et il peut aussi posséder son intérêt propre, quand ce ne serait qu'esthétique !

Prenons un exemple. On sait que les « pianistes », émetteurs clandestins en temps de guerre, courent toujours le risque d'être « retournés » par l'ennemi et d'être contraints d'émettre de faux messages destinés à tromper leur service de renseignements d'origine ; aussi doivent-ils, pendant l'émission, diffuser un certain signe personnel d'identification dont l'absence ou la modification permettra de faire connaître la contrainte qu'ils subissent. Imaginons que le code utilise trois signes (comme les signes Morse • — et / de la page 130), signes que nous noterons o, 1 et 2. Après chaque tranche de 30 signaux, il est convenu avec l'opérateur de considérer les trois derniers émis, soient (xyz) et d'envoyer ensuite trois signes $(x'\, y'\, z')$ définis de la façon suivante. x' n'est autre que $(y + z)$, calculé dans l'arithmétique F_3 (voir page 112) ; $y' = x + z$ et $z' = x + y$. Si les derniers signaux étaient (021), le pianiste doit donc émettre ensuite (012), puisque

$$\begin{cases} 2 + 1 = 0 \text{ (trois égale zéro dans } F_3), \\ 0 + 1 = 1, \quad 0 + 2 = 2. \end{cases}$$

L'opération qui consiste à passer du triplet (x, y, z) au triplet (x', y', z') est bien un calcul dans un espace vectoriel. On convient d'écrire les trois égalités :

$$\begin{cases} x' = 0 \times x + 1 \times y + 1 \times z \\ y' = 1 \times x + 0 \times y + 1 \times z \\ z' = 1 \times x + 1 \times y + 0 \times z \end{cases}$$

sous la forme symbolique unique :

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Elle relie trois objets (deux allongés et un carré) appelés *matrices*. D'une manière plus générale, on définit le produit d'une matrice rectangulaire M par une matrice-colonne X de la façon suivante :

$$Y = \begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = M \times X$$

si l'on a les relations linéaires :

$$\begin{cases} x' = ax + by + cz \\ y' = dx + ey + fz \\ z' = gx + hy + iz \\ t' = jx + ky + lz \end{cases}$$

Un tel produit est donc une sorte de codage d'un triplet (x, y, z) en un quadruplet (x', y', z', t'). Toute application linéaire d'un espace de dimension finie dont on a explicité une base dans un autre espace peut se traduire par un produit matriciel de ce genre, où (x, y, z) et $(x'\ y'\ z'\ t')$ sont les coordonnées du vecteur $\vec{p}$ et de son transformé $f(\vec{p})$.

Pour qu'un tel codage soit réversible, il est nécessaire (mais non suffisant) que la matrice qui le représente soit carrée. C'est le cas pour les relations

$$(x' = ox + y + 2z, y' = 3x + 0y + z, z' = x + 2y + 4z),$$

qui définissent le produit matriciel :

$$Y = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 0 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \times X$$

puisque l'on peut reconstituer le triplet (x, y, z) à partir du triplet (x', y', z') par les formules :

$$(x = -2x' + 0y' + z',\ y = -11x' - 2y' + 6z',\ z = 6x' + y' - 3z')$$

ou le produit matriciel équivalent :

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 & 0 & 1 \\ -11 & -2 & 6 \\ 6 & 1 & -3 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = B \times Y.$$

Les deux matrices A et B sont dites inverses l'une de l'autre.

Multiplier deux matrices

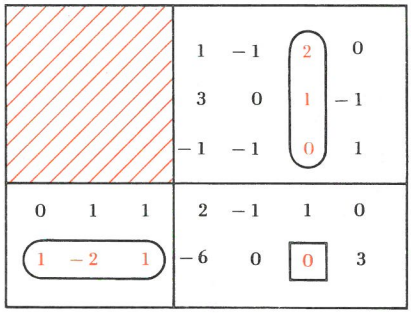
On peut même multiplier entre elles des matrices d'une forme plus générale encore. Pour définir le *produit* d'une matrice A possédant 2 lignes et 3 colonnes par une matrice B, il faut et il suffit que B ait 3 lignes (et, par exemple, 4 colonnes). Pour cela, on considère B comme la matrice d'une application linéaire (disons d'un codage) qui transforme (x, y, z, t) en $(x', y' z')$, et A comme transformant (x', y', z') en (x'', y''). Le passage direct de (x, y, z, t) à (x'', y'') est donc en quelque sorte le résultat d'un surcodage, mais on peut aussi le considérer comme un codage simple dont la matrice C est, par définition, le produit $A \times B$ dans cet ordre. Donnons un exemple :

$(x' = x - y + 2z + 0t,\ y' = 3x + 0y + z - t,$
$z' = -x - y + 0z + t)$ et $(x'' = 0x' + y' + z',$
$y'' = x' - 2y' + z')$ conduisent aux écritures matricielles :

$$Y = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 & 0 \\ 3 & 0 & 1 & -1 \\ -1 & -1 & 0 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = B \times X,$$

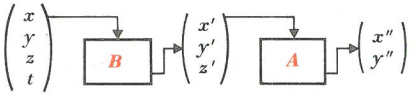
$$Z = \begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \times Y.$$

Un calcul élémentaire conduit directement aux égalités :
$(x'' = 2x - y + z + 0t,\ y'' = -6x - 2y + 0z + 3t)$
donc à l'écriture matricielle :

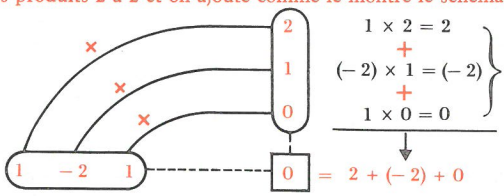


Produit de deux matrices.

Pour que le produit $A \times B$ soit défini, il faut que la figure hachurée soit un carré. Pour calculer un élément de $A \times B$ sans expliciter le processus :



On considère la ligne de A et la colonne de B qui se coupent à l'endroit choisi, on fait les produits 2 à 2 et on ajoute comme le montre le schéma :



$$Z = \begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 2 & -1 & 1 & 0 \\ -6 & -2 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = C \times X.$$

Par définition $C = A \times B$, ce que nous écrirons sous la forme :

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 & 0 \\ 3 & 0 & 1 & -1 \\ -1 & -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 1 & 0 \\ -6 & -2 & 0 & 3 \end{pmatrix}$$

(On peut effectuer directement ce calcul, sans passer par l'interprétation par les applications linéaires.) Ce produit de matrices (n lignes – p colonnes) par des matrices (p lignes – q colonnes) donne des matrices (n lignes – q colonnes). Si n est différent de q, le produit $B \times A$ ne peut être formé : la multiplication n'est donc pas commutative. Même si $n = q$,

les produits $A \times B$ et $B \times A$ sont distincts en général, puisque la première est carrée à n lignes, la seconde est carrée à p lignes. Enfin le cas très particulier $n = p = q$ ne donne généralement pas non plus $A \times B = B \times A$, puisque, par exemple :

$$A \times B = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 8 & -2 \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$B \times A = \begin{pmatrix} 8 & -2 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 10 & -20 \\ 5 & -10 \end{pmatrix} \neq A \times B.$$

Les matrices carrées

Limitons-nous désormais à des matrices carrées à n lignes. Leurs éléments appartiennent à un anneau, souvent à un corps. La somme de deux matrices est, par définition, la matrice dont les éléments sont les sommes de ceux des deux matrices placés en des endroits analogues : ainsi

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 0 & -3 \\ 0 & 4 & 2 \\ -2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 & -1 \\ 3 & 3 & 1 \\ -2 & 2 & 1 \end{pmatrix}.$$

Cette addition a un élément neutre (la matrice creuse qui n'a que des zéros). Muni de la multiplication et de l'addition, l'ensemble des matrices carrées est un anneau non commutatif (si $n > 1$) à diviseurs de zéro, comme le montre un exemple donné ci-dessus : $A \times B$ est creuse sans qu'aucune des deux ne le soit. Cet anneau possède un élément neutre pour la multiplication : c'est la matrice I dont tous les éléments sont nuls à l'exception de ceux de la « diagonale » qui sont égaux à 1 :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ (pour } n = 3).$$

C'est la matrice du codage « en clair », qui ne modifie rien.

Cet exemple est le plus simple d'une structure non commutative, même si elle est définie sur des anneaux ou corps commutatifs (comme $\mathbf{R}$). Son importance est très grande en physique notamment, où certains calculs de physique quantique, entre autres, exigent des produits non commutatifs et s'expriment facilement en termes de matrices (Heisenberg) [1].

1. Voir par exemple *l'Étrange histoire des quanta* de Banesh Hofmann, aux éditions du Seuil.

Le calcul matriciel possède des particularités très intéressantes. Bien que l'anneau des matrices carrées ne soit pas un corps (puisqu'il y existe des diviseurs de zéro), il contient un sous-ensemble qui est un groupe multiplicatif non commutatif (le groupe linéaire). Nous avons vu page 155 l'exemple de deux matrices inverses : leur produit était égal à I comme on peut facilement le voir en effectuant successivement le codage et le décodage de (x, y, z) en (x', y', z'). Certains sous-groupes du groupe linéaire sont très importants en géométrie ; le groupe orthogonal, par exemple, est constitué des matrices des applications linéaires qui généralisent les déplacements euclidiens (si on les considère comme des applications entre points géométriques et non entre vecteurs, ce sont celles qui conservent la distance de deux points). Cet anneau est également un vectoriel.

Les valeurs propres d'une application

Dans un exemple de codage matriciel (page 115), nous avons utilisé les transformations ($x' = 2y + 3z$, $y' = 2x + 3y$, $z' = 3x + 2z$) à éléments dans l'anneau $\mathbb{Z}_6$. La transformation pouvait s'écrire matriciellement :

$$Y = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 2 & 3 \\ 2 & 3 & 0 \\ 3 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \times X.$$

Elle possède la curieuse propriété suivante (ne pas oublier que l'on est dans l'anneau $\mathbb{Z}_6$, où $6 = 0$) : la matrice A est sa propre inverse ($A = A^{-1}$ ou encore $A^2 = A \times A = I$). D'autre part elle est invariante si l'on remplace chaque terme par le terme symétrique par rapport à la diagonale. Toute matrice P telle que l'on obtienne, par cette symétrie (appelée *transposition*), sa matrice inverse P^{-1} est justement l'une des matrices du groupe orthogonal dont nous avons mentionné l'existence.

Dans ce codage, il est intéressant de rechercher les triplets (x, y, z) qui restent invariants, c'est-à-dire tels que $x' = x$, $y' = y$ et $z' = z$. On trouve les triplets suivants ; (000) (030) (123) (153) (210) (240) (303) (333) (420) (450) (513) (543). D'autres sont transformés de manière encore relativement simple, par exemple en $(2x, 2y, 2z)$: c'est le cas de (000) (002) (004) (220) (222) (224) (440) (442) (444). Ces deux ensembles de triplets sont des espaces analogues à des espaces vectoriels ($\mathbb{Z}_6$ étant un anneau et non un corps,

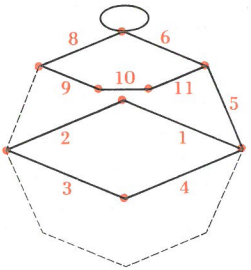
ces espaces sont appelés des modules). Tout nombre s tel qu'il existe un p non nul dont le transformé est $f(p) = s\,p$, est appelé *valeur propre* de l'application f et de la matrice qui le représente éventuellement. Notre matrice admet cinq valeurs propres : 1, 2, 3, 4 et 5. Elle transforme par exemple le triplet (032) en le triplet (034) qui est bien le triplet obtenu en multipliant par 5 chaque élément de (032) puisque $5 \times 0 = 0$, $5 \times 3 = 3$ (quinze égale deux fois six plus trois) et $5 \times 2 = 4$. (dix égale six plus quatre) ; 5 est donc une valeur propre.

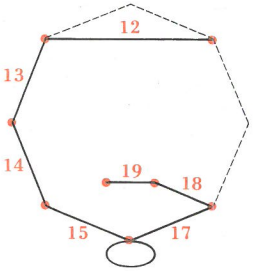
Les valeurs propres ont une grande importance, car elles correspondent à des propriétés de l'application linéaire indépendantes des bases choisies pour effectuer les calculs sur les coordonnées (on dit que ce sont des notions intrinsèques, invariantes dans tous changements de base).

Une des applications les plus importantes de la théorie des espaces vectoriels réside dans la notion d'*extension algébrique* d'un corps. Considérons un corps K et un élément x n'appartenant pas à K, mais que l'on suppose être racine d'une équation algébrique à coefficients dans K de la forme :

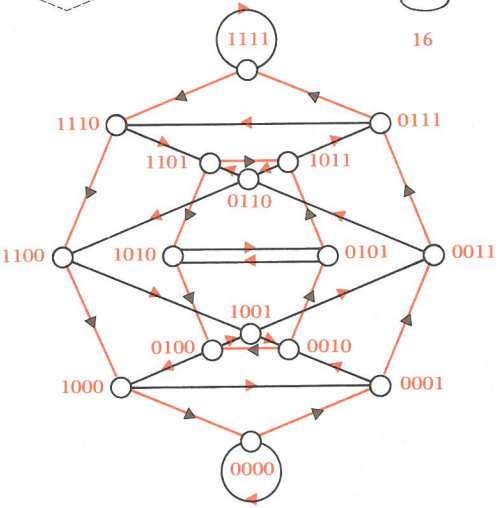
$$x^n = ax^{n-1} + bx^{n-2} + \ldots + jx + k.$$

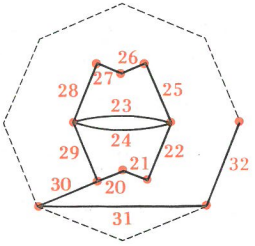
Posons par surcroît que le polynôme qu'annule x ne peut pas être décomposé en produit de polynômes de degrés inférieurs. Considérant x comme un nouvel élément de K, on est amené à calculer ses différentes puissances, x, x^2, ..., ainsi que des expressions telles que $1/x$, $1/x + 1$, etc. On peut montrer que toutes ces quantités peuvent s'écrire de manière unique sous la forme d'une combinaison linéaire de 1, x, x^2,..., x^{n-1} (par exemple $x^n = k + jx + \ldots + ax^{n-1}$). C'est dire que l'ensemble des expressions que l'on peut former à partir des éléments de K et de x en n'employant que des opérations rationnelles (additions, soustractions, multiplications et divisions) constitue un espace vectoriel de dimension n sur le corps K (et même un anneau), dont une base est $(1, x, x^2, \ldots, x^{n-1})$. On peut alors montrer que la véritable structure de cette extension algébrique de K, notée $K(x)$, est celle de corps ; K peut naturellement être assimilé à un sous-corps de $K(x)$. L'exemple le plus célèbre d'une telle extension est celui de $\mathbf{R}$ et de l'équation $i^2 = -1$. L'extension obtenue, sur-corps de $\mathbf{R}$ et espace vectoriel de dimension 2 sur $\mathbf{R}$, est le corps des *nombres complexes* $\mathbf{C} = \mathbf{R}(i)$; i, défini par $i^2 = -1$, est le germe unique dont la seule présence suffit pour engendrer l'ensemble de tous les nombres complexes.





16





00110011111011100000100

1010110100011

Un graphe de Good pour des quintuplets