Autour du concept de nombre

Fabien Besnard*

26 Janvier 2000

1 Introduction

Le concept de nombre peut paraître aujourd'hui intuitivement évident. Pourtant, en se penchant sur l'histoire des mathématiques, force est de constater que ce concept a extraordinairement évolué. En partant des entiers naturels, pour arriver aux quantités non-commutatives au comportement étrange, qui sont pourtant si présentes dans les mathématiques et la physique modernes, je vais tenter un rapide survol de cette évolution.

2 Les entiers naturels

Il est clair qu'il fut un temps où les hommes ignoraient tout du concept de nombre entier. Jusqu'à récemment[1], chez certaines peuplades, la suite des entiers s'arrêtait...à 2 ! (On comptait de la sorte : un, deux,...beaucoup)
Toutefois, ces hommes qu'on pourrait appeler "pré-numériques" ne sont pas démunis face aux situations quotidiennes. Ainsi, imaginons un homme de Cro-magnon, appelons-le Tik, chargé de veiller sur un troupeau de 8 moutons.
Tik ne sait pas compter, mais il aimerait bien savoir si tous ses moutons sont bien rentrés à l'enclos. Rien de plus facile ! Le premier jour, il fait sortir ses bêtes une par une, en faisant à chaque fois une encoche dans un os ou dans un bâton, par exemple. Dès lors, Tik n'aura qu'à faire rentrer ses moutons un par un, en faisant simultanément glisser son doigt d'une encoche à l'autre. Cette technique a été utilisée il y a au moins 30000 ans, comme en témoigne un radius de loup retrouvé en république Tchèque, comportant 45 encoches. Mais cette technique est plus qu'une astuce. Qu'a fait Tik ? Il a réalisé un appariement entre les moutons et les encoches, ce qu'en termes mathématiques on appelle une bijection, f, de l'ensemble M des moutons dans l'ensemble E des encoches. Rappelons qu'une bijection est une application qui possède les deux propriétés suivantes : elle envoie deux éléments distincts sur deux éléments distincts (elle est injective), et tout élément de l'ensemble d'arrivée est l'image d'au moins un élément de l'ensemble de départ (elle est surjective).

*besnard@math.jussieu.fr

[1]La plupart a aujourd'hui adopté le système de numération occidental

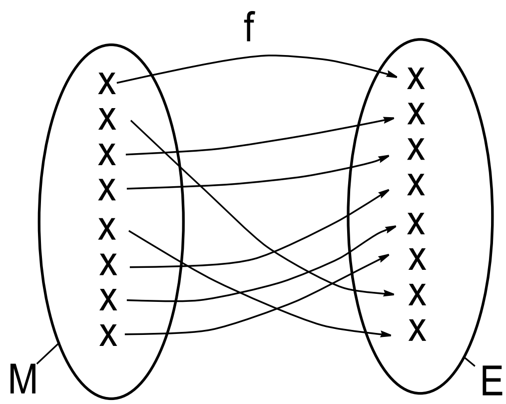

Fig.1. Une bijection

Mais Tik pourra bientôt se rendre compte que d'autres ensembles peuvent être mis en bijection avec E. Ces ensembles possèdent une propriété commune indépendante de la nature de leurs éléments. Cette propriété, on peut l'appeler "avoir 8 éléments", c'est ainsi que nait le concept abstrait du "nombre 8". Tik utilisera son bâton pour reconnaître les ensembles à 8 éléments : c'est un représentant particulier parmi ces ensembles, qui sert de référence. La suite des entiers naturels (0, 1, 2, etc...) n'est rien d'autre qu'une référence universelle arbitraire, apprise par coeur, qui a le même usage (en plus facile à transporter) qu'une infinité de bâtons de comptage !

Faisons un bond de plusieurs millénaires pour examiner la manière dont les mathématiciens de la fin du XIXe siècle et du début du XXe ont répondu à cette question. À cette époque, on a voulu fonder les mathématiques sur des bases rigoureuses[2]. Pour cela on a inventé la théorie des ensembles, ces entités semblant être à la base même de la pensée mathématique, comme j'ai essayé de l'illustrer avec mon exemple "préhistorique". Construire une théorie cohérente des ensembles soulève de nombreux problèmes, mais nous n'en parlerons pas, nous contentant de la théorie intuitive, qu'on enseigne au collège. Voyons donc comment on définit les entiers naturels à partir des ensembles, en reproduisant en quelque sorte l'idée de Tik.

Pour tout ensemble x, on note $s(x) = x \cup \{x\}$, qu'on appelle "successeur" de x. Notez bien que $x \cup \{x\} \neq x$! Par exemple, si x est l'ensemble $\{a;b\}$, $s(x) = \{a;b\} \cup \{\{a;b\}\} = \{a;b;\{a;b\}\}$.

Voyons ce que cela donne, en partant de l'ensemble vide : $s(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\}$, puis $s(s(\emptyset)) = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\}$, etc...On remarque que le vide contient 0 éléments, alors que $s(\emptyset)$ en contient un, et $s(s(\emptyset))$ deux, etc...On postule alors (c'est un axiome de la théorie des ensembles, qu'il existe un ensemble $\mathbf{N}$,

[2]C'était le cas aussi du temps d'Euclide, mais entre temps les critères de rigueurs ont été renforcés

qui contient $\emptyset$, et tous ses successeurs, et rien d'autre. Nous n'avons plus qu'à introduire les notations abrégées : $0 = \emptyset$, $1 = \{\emptyset\}$, etc...
C'est assez amusant de constater que tout est défini à partir du vide !

Mais entre l'Aurignacien et la fin du XIXe siècle, bien d'autres nombres que les entiers naturels ont été découverts ! Revenons un peu sur cette histoire.

3 Les fractions et les entiers relatifs

Les nécessités du calcul, notamment pour l'agriculture, l'architecture, ont conduit les égyptiens, les babyloniens, et d'autres encore, à utiliser des fractions. Cependant, celles-ci furent longues à être considérées comme des nombres à part entière. Euclide, par exemple, répugnait à les utiliser. À leur place, il avait construit une théorie subtile, celle des nombres "commensurables". Deux grandeurs A et B de même espèce (par exemple 2 longueurs, 2 aires, etc...), sont dites commensurables ssi il existe une autre grandeur C, de même espèce, et 2 entiers p et q, tels que :

$$A = pC$$

$$B = qC$$

Et ce n'est pas avant l'âge moderne (après la Renaissance), qu'on a pu écrire quelque chose comme :

$$\frac{A}{B} = \frac{p}{q}$$

Les calculs sur les fractions, de plus en plus considérées comme de vrais nombres, se sont tout de même répandus, notamment après Diophante, mais on a longtemps gardé l'idée que A et B devaient être des grandeurs de même nature. En particulier, Galilée n'a jamais été en mesure d'écrire la formule $v = \frac{d}{t}$! Nous allons bientôt voir comment les mathématiciens d'aujourd'hui définissent les fractions, mais comme ce n'est pas si évident, nous allons d'abord nous attaquer à une espèce de nombre qui de nos jours semble plus intuitive, et qui pourtant a mis encore plus de temps à rentrer dans les moeurs : les nombres négatifs.
En effet, même si les mathématiciens chinois s'en servaient au premier siècle après J.C., les nombres négatifs rencontrèrent encore des réticences chez certains de leurs confrères européens jusqu'au XVIIIe siècle. Peut-être parce qu'avant de passer aux négatifs il fallait d'abord découvrir le zéro, ce qui a pris du temps.
Mais même si aujourd'hui il semble facile de donner aux jeunes l'intuition des entiers relatifs, en se servant par exemple de l'image du thermomètre avec ses températures négatives, qui a contribué à populariser cette notion, la construction rigoureuse de ces nombres demande une attention un peu plus soutenue.
Connaissant $\mathbf{N}$ et l'addition, il n'est pas difficile, dans un premier temps, de définir les différences $a - b$, avec $a \geq b$. Un entier relatif sera alors simplement une "différence généralisée", où a pourra être plus petit que b. Comment rendre cette idée rigoureuse ?

À première vue, une différence est donnée par un couple d'entier $(a,b) \in \mathbf{N}\times\mathbf{N}$. Cependant, si $c \in \mathbf{N}$, $(a+c)-(b+c) = a-b$, donc les couples $(a+c, b+c)$ et (a,b) doivent représenter la même différence, or ce sont des couples différents. Pour nous sortir de ce mauvais pas, nous allons introduire la notion de "relation d'équivalence". Pour expliquer cette notion, nous allons partir d'un exemple complètement idiot ! Sur n'importe quel ensemble E, la relation "être égal à" est une relation d'équivalence, ce qui veut dire que pour tous $x, y, z \in E$, on a :

$$x = x$$

$$\text{si } x = y \text{ alors } y = x$$

$$\text{si } x = y \text{ et } y = z, \text{ alors } x = z$$

Mais vous connaissez sûrement une relation d'équivalence un peu moins stupide, qui est définie sur l'ensemble D des droites du plan (ou de l'espace) : la relation "être parallèle à". On a en effet, pour toutes droites $d, d', d'' \in D$:

$$d \parallel d$$

$$\text{si } d \parallel d' \text{ alors } d' \parallel d$$

$$\text{si } d \parallel d' \text{ et } d' \parallel d'', \text{ alors } d \parallel d''$$

Ce qui est intéressant avec une relation d'équivalence, c'est qu'elle découpe l'ensemble en "tranches", qu'on appelle des "classes d'équivalence". Par exemple, pour la relation $\parallel$, la classe d'équivalence de la droite d est l'ensemble de toutes les droites qui sont parallèles à d. Dans ce cas, la classe d'équivalence s'appelle "direction". Dans la figure 2, on a illustré ce phénomène sur un ensemble fini $\{a; b; c; d; e; f\}$, sur lequel on a défini la relation d'équivalence $\sim$, qui vérifie $a \sim b$, $c \sim d$, et $d \sim e$.

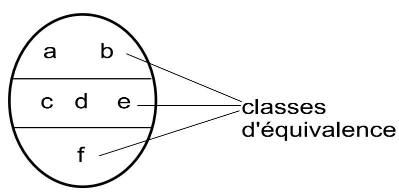

Fig.2. Partition d'un ensemble en classes d'équivalence. f est tout seul dans sa classe, car il n'est équivalent qu'à lui-même.

Revenons à nos couples d'entiers : définissons sur $\mathbf{N}\times\mathbf{N}$ la relation d'équivalence suivante :

$$(a,b) \sim (c,d) \text{ ssi } a+d = b+c \tag{1}$$

(On aurait pu avoir envie d'écrire $a-b = c-d$, mais cette soustraction n'est pas toujours possible, puisqu'on ne connaît pas encore les nombres négatifs !)

Vérifions que nous avons bien une relation d'équivalence : on a $(a, b) \sim (a, b)$, car $a + b = a + b$. Si $(a, b) \sim (c, d)$, alors $(c, d) \sim (a, b)$, car si $a + d = b + c$, on a aussi $c + b = d + a$. Je vous laisse le soin de vérifier la dernière propriété.
Ce que nous allons baptiser "entier relatif", ce sera une classe d'équivalence pour cette relation. Ça peut paraître bizarre, mais nous allons voir que nous avons bien "capté" la notion que nous voulions. En effet, si nous notons $\overline{(a, b)}$ la classe d'équivalence du couple (a, b), c'est-à-dire l'ensemble des couples qui sont équivalents au couple (a, b), on a par exemple :

$$\overline{(3, 2)} = \{(1, 0); (2, 1); (3, 2); (4, 3); \ldots\}$$

(ce sont tous les couples (a, b) tels que la différence $a - b$ vaut 1) ou encore :

$$\overline{(2, 3)} = \{(0, 1); (1, 2); (2, 3); \ldots\}$$

(ce sont tous les couples (a, b) tels que $b - a = 1$)

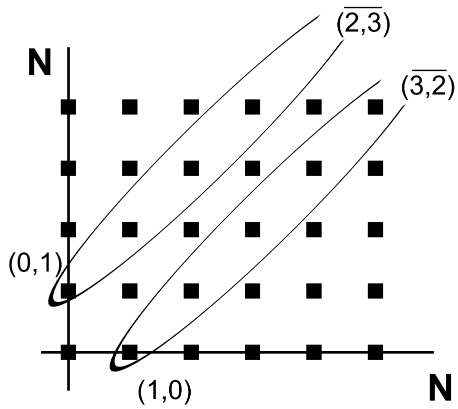

Fig.3. Les classes d'équivalences $\overline{(3, 2)}$ et $\overline{(2, 3)}$.

On remarque immédiatement que dans $\overline{(3, 2)}$ il y a un élément plus sympathique : c'est $(1, 0)$, et dans $\overline{(2, 3)}$, il y a $(0, 1)$. Donc $\overline{(3, 2)} = \overline{(1, 0)}$, et $\overline{(2, 3)} = \overline{(0, 1)}$. De même, si $a \geq b$, $\overline{(a, b)} = \overline{(a - b, 0)}$, et si $a \leq b$, $\overline{(a, b)} = \overline{(0, b - a)}$. Pour retrouver les entiers relatifs que nous connaissons et aimons, il n'y aura plus qu'une étape à franchir : introduire la notation $\overline{(c, 0)} = c$, et $\overline{(0, c)} = -c$.
Mais avant toute chose, il nous faut dire comment on additionne ou multiplie deux entiers relatifs, sachant comment l'on fait avec des entiers naturels. Il suffit pour cela de s'inspirer des formules pour les différences d'entiers naturels :

$$(a - b) + (c - d) = (a + c) - (b + d)$$

$$(a - b) \times (c - d) = (ac + bd) - (ad + bc)$$

Ce qui, traduit pour les couples, donne :

$$(a,b)+(c,d)=(a+c,b+d) \tag{2}$$
$$(a,b)\times(c,d)=(ac+bd,ad+bc) \tag{3}$$

Mais pour passer à $\mathbf{Z}$, encore faut-il passer des couples aux classes, c'est-à-dire mettre une barre sur la tête des deux formules précédentes :

$$\overline{(a,b)}+\overline{(c,d)}=\overline{(a+c,b+d)} \tag{4}$$
$$\overline{(a,b)}\times\overline{(c,d)}=\overline{(ac+bd,ad+bc)} \tag{5}$$

Ceci est possible à une condition : que dans les formules (2) et (3) les résultats ne changent pas lorsqu'on remplace un couple par un couple équivalent. En effet, on a par exemple $(3,1)\sim(4,2)$, donc $\overline{(3,1)}=\overline{(4,2)}$. Par conséquent, les formules (4) et (5) ne doivent pas changer si on prend $a=3$ et $b=1$, ou $a=4$ et $b=2$.

On peut vérifier facilement que cette condition est vérifiée.

C'est maintenant que nous pouvons pleinement justifier la notation $\overline{(c,0)}=c$. En effet, rien ne distingue du point de vue de l'addition et de la multiplication l'entier c de la classe $\overline{(c,0)}$, puisqu'on a :

$$\overline{(a,0)}+\overline{(b,0)}=\overline{(a+b,0)} \tag{6}$$
$$\overline{(a,0)}\times\overline{(b,0)}=\overline{(ab,0)} \tag{7}$$

Nous pouvons donc identifier $\mathbf{N}$ avec l'ensemble des classes du type $\overline{(c,0)}$, et considérer que $\mathbf{N}\subset\mathbf{Z}$.

Passons maintenant à $\mathbf{Q}$. La méthode est très similaire. Pour se donner une fraction a/b, il faut naturellement se donner un couple d'entiers $(a,b)\in\mathbf{Z}\times\mathbf{Z}^*$ avec $b\neq 0$. Mais si $ad=bc$, les fractions a/b et c/d sont égales, on va donc identifier les couples correspondants, en introduisant la relation d'équivalence :

$$(a,b)\sim(c,d)\Leftrightarrow ad=bc \tag{8}$$

(remarquer que c'est exactement la formule (1) à ceci près que la multiplication remplace l'addition)

On s'inspirant des formules

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd},\quad \frac{a}{b}\frac{c}{d}=\frac{ac}{bd} \tag{9}$$

On définit aussi une addition et une multiplication pour les couples :

$$(a,b)+(c,d)=(ad+bc,bd),\quad (a,b)(c,d)=(ac,bd) \tag{10}$$

On vérifie que ces deux lois sont compatibles avec la relation d'équivalence (i.e. elles ne changent pas si on remplace un couple par un couple équivalent), et on appelle "fraction a/b" la classe d'équivalence du couple a/b. On notera bien sûr $\mathbf{Q}$ l'ensemble des fractions. De même qu'on a identifié $\mathbf{N}$ avec un sous-ensemble de $\mathbf{Z}$, on identifie $\mathbf{Z}$ avec l'ensemble des fractions du type $a/1$, et on se permet d'écrire $\mathbf{Z}\subset\mathbf{Q}$.

4 Les nombres réels

Les nombres que nous avons construits jusqu'ici permettent de résoudre des équations, comme $2x = 3$ ou $x + 8 = 5$ qui sinon seraient impossibles. Mais depuis la plus haute antiquité, d'autres équations ont eveillé la curiosité des mathématiciens : celles du second degré.
Or, les pythagoriciens avaient découvert que l'équation $x^2 = 2$ n'a pas de solution dans $\mathbf{Q}$ (tout cela étant dit en termes modernes, bien sûr). Rappelons la démonstration, qui est très simple. On suppose qu'une solution $x \in \mathbf{Q}$ existe, et on écrit x sous forme de fraction irréductible, c'est-à-dire $x = p/q$, avec p et q deux entiers premiers entre eux. On a donc $p^2 = 2q^2$, ce qui montre que p^2 est pair. Mais comme le carré d'un nombre impair est impair, p ne saurait être impair, il est donc pair. On peut donc écrire $p = 2p'$, avec p' entier. Mais alors on a $(2p')^2 = 2q^2$, ce qui donne $2p'^2 = q^2$, et par le même raisonnement, on voit que q est pair. Mais on avait supposé que p et q étaient premiers entre eux, c'est donc absurde.
Ceci a profondément troublé les pythagoriciens, car ils pensaient que tous les nombres étaient rationnels. On dit qu'il se transmettaient le secret de la preuve de l'irrationalité de $\sqrt{2}$ uniquement entre initiés !
Ce qui est en effet troublant, ce que l'on peut difficilement décréter que l'équation $x^2 = 2$ n'admet pas de solution, puisque c'est l'équation que doit vérifier la longueur de la diagonale d'un carré de côté 1 !

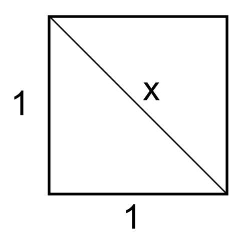

Fig.4. $x^2 = 2$

On voit ici que c'est la géométrie qui va pousser les mathématiciens à enrichir leur bestiaire de nombres. Mais ceci va les conduire à une nouvelle abstraction : la notion de limite.
En effet, on peut facilement construire une suite (u_n) de nombres rationnels, qui s'approche de plus en plus de x, sans jamais l'atteindre, par la méthode des intervalles emboîtés.
Comme $1^2 \leq 2 \leq 2^2$, on en déduit que $x \in [1, 2]$. On commence donc par poser $u_1 = 1.5$, qui sera notre première approximation de x. Mais comme $u_1^2 \geq 2$, on sait que $x \in [1, u_1]$, et on pose $u_2 = \dfrac{1 + u_1}{2} = 1.25$. Cette fois $u_2^2 \leq 2$, donc $x \in [u_2, u_1]$, et on pose $u_3 = \dfrac{u_1 + u_2}{2}$, etc...On peut voir facilement que les

termes de la suite se rapprochent de plus en plus les uns des autres (vous pouvez vous amusez à vérifier que pour tout n et pour tout p, $|u_{n+p} - u_n| \leq 2^{-n}$). On dit que c'est une suite de Cauchy. En décrétant qu'il existe une limite unique, notée $\sqrt{2}$, on crée une nouvelle sorte de nombre, obtenu comme limite d'une suite de rationnels, ou si l'on veut, par approximations successives. Mais a-t-on le droit de décréter cela ? Rappelons-nous quelle était la situation lorsque nous avons construit les entiers relatifs : nous voulions résoudre une équation du type $x + a = b$, avec $a \geq b$. Nous aurions pu aussi "décréter", c'est-à-dire poser des axiomes, que de tels nombres existaient, et qu'on pouvait leur appliquer les règles usuelles de l'algèbre, en ajoutant simplement la "règle des signes" ($- \times - = +$, etc...). En fait, c'est comme cela qu'on fait dans l'enseignement scolaire. Mais on pourrait se demander si tout cela a bien un sens ? La construction de **Z** à partir de **N** que nous avons donnée montre que si **N** a un sens, alors **Z** aussi. De la même manière, nous pouvons nous "rassurer" en construisant à partir de **Q** un ensemble **R**, constitué de "classes d'équivalence de suites de Cauchy", et vérifiant toutes les propriétés que l'on est en droit d'attendre des nombres réels.

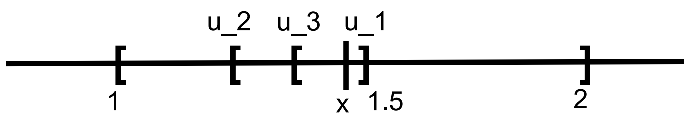

Fig.5. Approximation de $\sqrt{2}$.

Je vous épargnerai cette construction, mais à la place, je vais résumer la situation jusqu'à maintenant. Au début nous avons dit ce qu'était un entier, et nous avons admis qu'il existe un ensemble, **N**, les contenant tous, et ne contenant qu'eux[3]. Puis, nous avons construit un ensemble **Z** et des règles d'addition et de multiplication se comportant exactement comme il faut[4]. Nous avons ensuite évoqué la construction de **Q** sur le même principe. Jusque là, on pouvait hausser les épaules devant tant de sophistication inutile, alors que "l'on sait bien ce que sont les nombres rationnels". Toutefois, on voit que plus l'on s'éloigne de l' expérience sensible, par exemple par l'introduction des nombres réels (pourquoi y a-t-il un élément, de surcroît unique, dans l'intersection de tous les intervalles emboîtés ? Bonne question !), plus on a besoin de s'appuyer sur des constructions rigoureuses, de plus en plus abstraites. Mais la suite de l'histoire devrait convaincre même les plus sceptiques...

[3] C'est le maillon faible de la chaîne, mais ceci est une autre histoire !

[4] En fait, nous n'avons pas défini l'addition et la multiplication sur **N**, mais ceci n'est pas difficile.

5 Les nombres complexes

Il se trouve qu'en plus de la découverte des nombres irrationnels, l'étude des équations polynomiales, du type :

$$a_n x^n + \ldots + a_1 x + a_0 = 0 \tag{11}$$

a permis de mettre en évidence un autre type de nombres : les nombres complexes[5].
En effet, les algébristes italiens du XVIe siècle ont trouvé une méthode pour résoudre les équations du 3e degré. On élimine d'abord le terme du 2e degré (c'est toujours possible, par un changement de variable approprié). L'équation se présente donc sous la forme :

$$x^3 = px + q \tag{12}$$

On voit facilement que si u et v vérifient :

$$uv = \frac{p}{3} \tag{13}$$

$$u^3 + v^3 = q \tag{14}$$

alors $x = u + v$ est solution de (12). Mais grâce à (13) et (14) on connaît la somme et le produit de u^3 et v^3. On peut donc trouver ces 2 derniers nombres comme solutions de l'équation du second degré :

$$X^2 - qX + \frac{p^3}{27} = 0 \tag{15}$$

Donc $u^3 = \frac{q + \sqrt{\Delta}}{2}$ et $v^3 = \frac{q - \sqrt{\Delta}}{2}$, où $\Delta = q^2 - \frac{4p^3}{27}$ est le discriminant de l'équation (15). Donc :

$$x = u + v = \sqrt[3]{\frac{q + \sqrt{\Delta}}{2}} + \sqrt[3]{\frac{q - \sqrt{\Delta}}{2}} \tag{16}$$

Ceci n'a de sens que si $\Delta \geq 0$. Toutefois, si on prend par exemple $p = 6$, $q = -4$, on a $\Delta = -16$. Et en faisant comme si cela avait un sens, on trouve grâce à (15) : $u^3 = \frac{-4 + \sqrt{-16}}{2} = -2 + 2\sqrt{-1}$, et $v^3 = -2 - 2\sqrt{-1}$. Pour continuer ces calculs délirants, il faut prendre "la racine cubique" de ces 2 quantités ! Or, on constate en utilisant la formule $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ que $(1 + \sqrt{-1})^3 = -2 + 2\sqrt{1} = u^3$, et $(1 - \sqrt{-1})^3 = -2 - 2\sqrt{-1} = v^3$. On voit que si on pose :

$$x = u + v = 1 + \sqrt{-1} + 1 - \sqrt{-1} = 2$$

[5] Il ne faudrait pas déduire de ce qui précède que les tous les nombres réels sont solutions d'équations du type (11) à coefficients rationnels. Ce n'est pas le cas de π, par exemple, comme l'a démontré Von Lindemann en 1882

on a bien une solution de l'équation $x^3 = 6x - 4$! Donc, en utilisant ces quantités "imaginaires" que sont les racines de nombres négatifs dans les formules, on voit qu'elles s'éliminent mutuellement, et qu'on trouve une solution de l'équation. Les mathématiciens ont donc continué à utiliser, bien que souvent à contre-coeur, les nombres imaginaires afin de trouver des solutions aux équations polynomiales. Ce faisant, ils n'ont jamais rencontré aucune contradiction : tout se passe comme si on pouvait ajouter aux nombres réels un nombre imaginaire $i = \sqrt{-1}$, vérifiant $i^2 = -1$, et considérer l'ensemble $\mathbf{C}$ des nombres complexes $a + ib$, $a, b \in \mathbf{R}$, sur lesquels on calcule exactement comme on a l'habitude. On peut même montrer que tout polynôme de degré n possède exactement n racines dans $\mathbf{C}$, en comptant éventuellement les racines multiples (c'est le théorème de D'Alembert-Gauss)[6].
Mais il est très facile de justifier tous ces calculs : en définissant sur les couples de réels les lois suivantes :

$$(a, b) + (a', b') = (a + a', b + b') \tag{17}$$
$$(a, b)(a', b') = (aa' - bb', ab' + ba') \tag{18}$$

On peut alors vérifier que l'addition et la multiplication que nous venons de définir possèdent toutes les propriétés usuelles : commutativité, associativité, distributivité de la multiplication par rapport à l'addition, etc...En outre, les couples de la forme $(a, 0)$ se comportent exactement comme des réels : $(a, 0) + (a', 0) = (a+a', 0)$, $(a, 0)(a', 0) = (aa', 0)$. On remarque aussi qu'on a $(1, 0)(a, b) = (a, b)$, pour tout couple (a, b), et $(0, 1)^2 = (0, 1)(0, 1) = (-1, 0)$. Enfin, on a :

$$(a, b) = (a, 0) + (0, 1)(b, 0)$$

Il suffit donc maintenant d'identifier le couple $(a, 0)$ avec le réel a, et de poser $i = (0, 1)$ pour retrouver la notation :

$$(a, b) = a + ib$$

C'est exactement ce qu'a fait Gauss en introduisant la représentation du plan complexe. Dans le langage d'aujourd'hui on dirait simplement qu'un couple de réels n'est rien d'autre qu'un vecteur du plan, et que les lois $(a, b) + (a', b') = (a + a', b + b')$ et $(a, 0)(a', b') = (aa', ab')$ correspondent exactement à l'addition des vecteurs et à la multiplication d'un vecteur par un scalaire.
Les nombres complexes ont permis la création de théories entièrement nouvelles, d'une importance considérable, telle l'analyse complexe. Ils se sont de plus révélés de puissants outils dans presque toutes les branches des mathématiques, et par ricochet dans certaines parties importantes de la physique (certains d'entre vous s'en sont peut-être déjà servi en électricité). Mais on pourait croire que ce ne sont là que d'habiles raccourcis, dont la physique saurait éventuellement se passer. C'est peut-être vrai pour l'électricité, mais ça ne l'est plus du tout pour la physique quantique : cette théorie utilise de manière fondamentale les nombres complexes.

[6]Toutefois, pour les équations de degré 5 ou plus, il n'existe pas de formule telle que (16) qui donne les racines, comme l'a montré Galois.

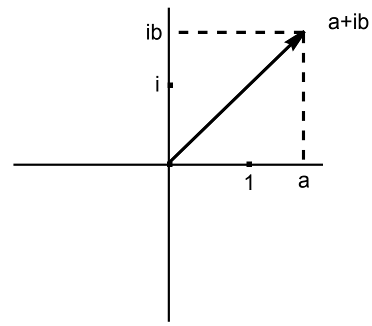

Fig.6. Le complexe $z = a + ib$ vu comme vecteur du plan. Les vecteurs horizontaux sont les nombres réels.

6 Au-delà des nombres complexes

Après avoir introduit une loi de multiplication sur les vecteurs du plan, il paraît naturel de chercher à généraliser les nombres complexes en définissant une loi de multiplication sur les vecteurs de l'espace. C'est ce qu'a cherché le mathématicien W. R. Hamilton pendant des années. Précisons la question : si $u = (a, b, c)$ et $v = (a', b', c')$ sont 2 vecteurs de l'espace, comment leur associer un 3e vecteur, noté uv, de sorte que l'on ait les propriétés usuelles d'une multiplication ? Mais qu'entendons-nous exactement par "propriétés usuelles" ? Nous voulons certainement que la mutliplication soit distributive par rapport à l'addition : $(u + v)(u' + v') = uu' + uv' + vu' + vv'$. Nous voulons aussi qu'elle soit associative : $(uv)w = u(vw)$. Il est aussi très important de pouvoir retrouver les nombres que nous connaissons déjà : tout comme les réels s'identifient aux nombres complexes $(a, 0)$ dont la 2e composante est nulle, il est naturel de demander que les complexes (a, b) s'identifient aux triplets $(a, b, 0)$. Enfin, tout comme chaque complexe non-nul z possède un inverse[7] z^{-1}, tel que $zz^{-1} = z^{-1}z = 1$, nous voulons que tout triplet $u \neq (0, 0, 0)$ possède un inverse. Hamilton a fini par réaliser qu'il n'existe pas de telle multiplication sur les vecteurs de l'espace. En revanche, en 1843, il a trouvé une manière de multiplier non plus des triplets, mais des quadruplets $q = (a, b, c, d)$, vérifiant toutes les propriétés précédentes ! Ces nouveaux nombres, qui généralisent les nombres complexes, sont appelés "quaternions", et leur ensemble se note $\mathbf{H}$. Regardons-les d'un peu plus près : posons $1 = (1, 0, 0, 0)$ et $i = (0, 1, 0, 0)$, pour identifier les quaternions de la forme $(a, b, 0, 0)$ avec les nombres complexes. Posons également $j = (0, 0, 1, 0)$, et $k = (0, 0, 0, 1)$. On a la table de mutliplication suivante :

$$i^2 = j^2 = k^2 = -1$$

$$ij = k, jk = i, ki = j$$

[7] Si $z = (a, b)$, $z^{-1} = (\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2})$

$$ji = -k, kj = -i, ik = -j$$

Comme tout quaternion q s'écrit $a + i.b + j.c + k.d$, tous les calculs se déduisent de cette table et des lois d'associativité et de distributivité. Tout comme le nombre complexe $z = a + i.b$ possède un conjugué $\bar{z} = a - i.b$, et une norme $|z| = \sqrt{a^2 + b^2}$, tels que $z\bar{z} = |z|^2$, tout quaternion $q = a + i.b + j.c + k.d$ possède un conjugué $\bar{q} = a - i.b - j.c - k.d$ et une norme $|q| = \sqrt{a^2 + b^2 + c^2 + d^2}$, tels que $q\bar{q} = \bar{q}q = |q|^2$. Comme tout quaternion q non-nul a une norme non-nulle, on en déduit que $q(\frac{\bar{q}}{|q|^2}) = 1$, donc $\frac{\bar{q}}{|q|^2}$ est l'inverse de q. Mais bien sûr j'ai gardé le meilleur pour la fin : on voit dans la table que $ij = k$ et $ji = -k$. Donc la multiplication des quaternions n'est pas commutative : pour 2 quaternions q et q', en général on a $qq' \neq q'q$!

Avec les quaternions nous devons aussi abandonner toute représentation visuelle : comme ils possèdent 4 coordonnées, il faudrait un espace à 4 dimensions pour les représenter !

Contrairement à **Z**, **Q**, **R** ou **C**, **H** n'a pas été introduit pour résoudre certains problèmes, mais, pourrait-on dire, juste "pour le fun". Cependant, il ne s'est pas écoulé un siècle avant que ceux-ci n'apparaissent d'eux-même en physique quantique, dans l'étude du spin de l'électron.

Peut-on continuer à ajouter des coordonnées pour trouver des nombres généralisant les quaternions ? La réponse est à la fois oui et non. Si on veut garder toutes les propriétés des quaternions (distributivité, associativité, existence d'inverse), la réponse est non : ça s'arrête là[8]. Mais si on est prêt à abandonner d'autres propriétés, alors on peut trouver quantité d'autres "nombres généralisés". Nous nous contenterons d'évoquer ce qu'on appelle "les matrices", dont vous avez peut-être déjà vaguement entendu parlé.

Les matrices ne sont rien d'autre que des tableaux de nombres. Lorsque l'on veut résoudre un système d'équations linéaires, comme par exemple :

$$\left\{\begin{array}{ccccccc} 3x_1 & + & 2x_2 & - & \sqrt{2}x_3 & = & 5 \\ 2x_1 & + & x_2 & + & 3x_3 & = & 0 \\ x_1 & - & x_2 & & & = & 2 \end{array}\right. \tag{19}$$

on lui associe la matrice 3×3 des coefficients :

$$M = \begin{pmatrix} 3 & 2 & -\sqrt{2} \\ 2 & 1 & 3 \\ 1 & -1 & 0 \end{pmatrix} \tag{20}$$

le vecteur des inconnues (qui n'est rien d'autre qu'une matrice 3×1) :

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{21}$$

[8]du moins en dimension finie !

et le vecteur des constantes :

$$Y = \begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix} \tag{22}$$

Le système (19) s'écrit alors sous forme matricielle $MX = Y$, c'est-à-dire :

$$\begin{pmatrix} 3 & 2 & -\sqrt{2} \\ 2 & 1 & 3 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix} \tag{23}$$

La signification de cette notation est la suivante : on multiplie la matrice M par le vecteur X et le résultat est le vecteur Y. Pour multiplier M par X, on prend le premier coefficient de la première ligne de M et on le multiplie par le 1er coefficient de X, puis le deuxième coefficient de la 1ere ligne de M, multiplié par le 2e coefficient de X, et enfin le 3e coefficient de la 1ere ligne de M multiplié par le 3e coefficient de X. On additionne le tout, cela donne : $3x_1 + 2x_2 - \sqrt{2}x_3$, qui est le premier coefficient du vecteur produit MX. On fait pareil avec la 2e ligne de M, et on obtient $2x_1 + x_2 + 3x_3$, qui est le 2e coefficient de MX. Enfin on fait la même chose avec la 3e ligne et on obtient le 3e coefficient. L'équation $MX = Y$ se traduit donc exactement par le système (19). On peut dire les choses autrement : on prend la première ligne de M, ce sont 3 nombres, qu'on considère comme les coordonnées d'un vecteur (un vecteur ligne), et on fait le produit scalaire de ce vecteur avec le vecteur (le vecteur colonne) X : on obtient un nombre qui est la 1ere composante du vecteur (colonne) MX.
Mais rien ne justifie qu'on se limite aux matrices 3×3 et aux vecteurs à 3 composantes. Voici la recette générale pour multiplier une matrice $m \times n$ (m lignes, n colonnes) que nous appelons A, et un vecteur colonne à n lignes (notez qu'il faut que la matrice ait autant de colonnes que le vecteur a de lignes) que nous appelons V : le résultat est un vecteur colonne à m lignes obtenu de la façon suivante :

> "le produit scalaire de la i-ème ligne de M avec V donne la i-ème composante de W"

Mais si vous savez multiplier une matrice par un vecteur, alors vous savez multiplier 2 matrices entre elles : il suffit de considérer une matrice comme la juxtaposition des vecteurs formés par ses colonnes. À titre d'exemple, multiplions la matrice M par la matrice suivante :

$$N = \begin{pmatrix} 1 & 5 \\ -1 & -7 \\ 1 & 0 \end{pmatrix}$$

La matrice N est composée des 2 colonnes :

$$N_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 5 \\ -7 \\ 0 \end{pmatrix}$$

Le produit MN est alors la matrice composée des 2 vecteurs colonnes MN_1 et MN_2, c'est-à-dire :

$$MN = \begin{pmatrix} 3\times 1 + 2\times(-1) + (-\sqrt{2})\times 1 & 3\times 5 + 2\times(-7) + (-\sqrt{2})\times 0 \\ 2\times 1 + 1\times(-1) + 3\times 1 & 2\times 5 + 1\times(-7) + 3\times 0 \\ 1\times 1 + (-1)\times(-1) + 0\times 1 & 1\times 5 + (-1)\times(-7) + 0\times 0 \end{pmatrix}$$

$$MN = \begin{pmatrix} 1-\sqrt{2} & 1 \\ 4 & 3 \\ 2 & 12 \end{pmatrix}$$

En revanche on ne peut pas faire le produit NM car la matrice N n'a que 2 colonnes alors que M a 3 lignes. Un cas particulier est celui des matrices carrées : on peut toujours mutliplier 2 matrices carrées entre elles. Mais on peut aussi additioner des matrices carrées, coefficient par coefficient. Par exemple :

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

On note $M_n(\mathbf{R})$ l'ensemble des matrices carrées $n\times n$ à coefficients dans $\mathbf{R}$, et $M_n(\mathbf{C})$ l'ensemble des matrices de même format mais à coefficient dans $\mathbf{C}$. Dans la suite on va examiner $M_n(\mathbf{R})$, mais tout ce qu'on va dire serait aussi valable pour $M_n(\mathbf{C})$. Pour fixer encore plus les idées on va prendre $n = 2$.
Comme nous avons une addition et une multiplication sur $M_2(\mathbf{R})$, nous pouvons nous demander si ces opérations vérifient les propriétés dont nous avons l'habitude. En ce qui concerne l'addition, tout se passe très bien, elle est associative, commutative, il y a un élément neutre (la matrice nulle, dont tous les coefficients sont nuls), et toute matrice M admet une matrice opposée $-M$, dont les coefficients sont les opposés de ceux de M. Pour la multiplication, cela sort un peu plus de l'ordinaire. Elle est toujours associative : si M, N et P sont 3 matrices, alors $M(NP) = (MN)P$. De plus elle est distributive par rapport à l'addition : $(M+N)(P+Q) = MP+MQ+NP+NQ$, donc on peut développer les expressions comme on en a l'habitude, à condition toutefois de faire attention à l'ordre des facteurs, car elle n'est pas commutative. En effet, on a par exemple :

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 7 & 4 \end{pmatrix}$$

alors que

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 5 \end{pmatrix}$$

La matrice notée I_2, dont les coefficients sont les suivants :

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

est l'élément neutre pour la multiplication, c'est-à-dire que pour toute matrice M, on a : $MI_2 = I_2M = M$.

La dernière chose à faire pour pouvoir considérer ces matrices comme des "nombres généralisés" est de reconnaitre parmi elles celles qui joueront le rôle de nos "anciens nombres". Il s'agira des matrices de la forme :

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

qu'on écrira aussi aI_2, la multiplication d'un réel a par une matrice M étant simplement la multiplication de tous les coefficients de M par a. On peut vérifier immédiatement que :

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a+b & 0 \\ 0 & a+b \end{pmatrix}$$

et

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} ab & 0 \\ 0 & ab \end{pmatrix}$$

Ainsi les calculs sur ces matrices reviennent exactement aux mêmes calculs sur des réels. On peut donc identifier le réel a et la matrice aI_2. Cependant, contrairement aux généralisations des nombres réels que nous avons déjà rencontrées (complexes, quaternions), celle-ci ne possède pas une propriété importante : l'existence d'inverse. En effet, la matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ne possède d'inverse, c'est-à-dire de matrice A^{-1} telle que $AA^{-1} = A^{-1}A = I_2$, que dans le cas où $ad - bc \neq 0$.
Avec, pour tout n, les matrices carrées $n \times n$, nous disposons d'une infinité de généralisations des nombres réels (ou complexes), mais cela ne s'arrête pas là. Pour aller plus loin on peut vouloir considérer des matrices infinies, avec une infinité de lignes et de colonnes. Cela peut se faire à condition de prendre certaines précautions. Mais le plus étrange, c'est que ce sont ces matrices infinies, en quelque sorte la généralisation "ultime" de la notion de nombre, qui ont, les premières, fait irruption en physique ! Et pas seulement de manière marginale : en mécanique quantique ce sont ces matrices infinies, qu'on appelle "opérateurs", qui remplacent les nombres usuels de la mécanique classique. Par exemple, en mécanique quantique la position x et la quantité de mouvement p d'une particule ne sont pas des nombres réels comme en mécanique classique, mais des opérateurs, $\hat{x}$ et $\hat{p}$, qui vérifient :

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar I \tag{24}$$

où $\hbar$ est un nombre réel très petit (c'est la constante de Planck divisée par 2π), i est le nombre complexe $\sqrt{-1}$, et I est l'opérateur "identité", qui généralise la matrice I_2 aux dimensions infinies. De la relation (24) découle le fameux principe d'incertitude de Heisenberg. C'est l'extrême petitesse de la constante

$\hbar$ qui nous fait croire, dans le monde macroscopique que nous connaissons, que les grandeurs physiques sont des nombres “normaux”, qui commutent.
Récemment, les algèbres non-commutatives telles que celles qui interviennent en physique quantique ont fait l'objet d'une “géométrisation”. Descartes avait jeté un pont entre l'algèbre (commutative) et la géométrie par sa méthode des coordonnées. Ce qu'a fait Connes avec sa géométrie non-commutative, c'est (en gros) considérer que les nombres non-commutatifs tels que $\hat{x}$ et $\hat{p}$ sont des coordonnées sur un espace “virtuel” non-commutatif ! Le but n'est rien moins que de réconcilier la relativité générale, qui est une théorie géométrique de l'espace-temps, et la théorie quantique des champs, en rendant l'espace-temps non-commutatif.

7 conclusion

J'espère avoir réussi à montrer que l'abstraction toujours plus grande vers laquelle ont évolué les mathématiques, et en particulier la notion de nombre, n'est pas due à une volonté malsaine des mathématiciens, mais à un soucis de rigueur et de généralité, qui a permis de se débarasser des idées reçues et des fausses évidences. Ce faisant, ils ont découvert des objets étranges, sur lesquels l'expérience sensible n'a pas de prise, et qui pourtant sont apparus naturellement dans l'étude des phénomènes microscopiques, montrant ainsi que la distinction entre physique et mathématiques est parfois quelque peu arbitraire.
Au fait, si j'ai appelé mon homme de cro-magnon “Tik”, c'est parce qu'il s'agirait, selon le linguiste Merrit Ruhlen, d'une racine ancienne et universelle : on la retrouve dans toutes les langues de la Terre, avec des sens toujours liés aux doigts ou aux nombres[9]. Ainsi, le mot le plus ancien a peut-être été prononcé par un de nos lointain ancêtres qui comptait sur ses doigts !

[9] Ainsi “digital” qui en Anglais signifie “numérique”, et en Français “relatif aux doigts”, contient cette racine linguistique

8 Repères chronologiques, références

- environ -30000 ans : Technique de l'entaille
- Ve siècle av. J.C. : Pythagore, irrationalité de $\sqrt{2}$
- IIIe siècle av. J.C. : Euclide fonde la méthode axiomatique, Archimède donne $3 + 10/70 \leq \pi \leq 3 + 10/71$
- Ier siècle : Les chinois connaissent les nombres < 0
- IIIe siècle : Diophante calcule sur des fractions, en dépit d'Euclide
- IVe siècle : Le zéro est inventé en Inde
- 1484 : N. Chuquet introduit les nombres < 0 en Europe
- 1545-1560 : Introduction des nombres imaginaire par Cardano et Bombelli
- 1637 : Descartes, Géométrie analytique : un pont se crée entre algèbre et géométrie
- 1760 : Lambert démontre que π est irrationnel
- 1800 : Gauss introduit le plan complexe
- 1843 : Hamilton invente les quaternions
- 1858 : Cayley, calcul matriciel
- 1830-1872 : Cauchy, Dedekind, définition précise des nombres réels.
- 1882 : Von Lindemann démontre la transcendance de π
- Fin XIXe-début XXe : Cantor, Peano, Zermelo, etc... : théorie des ensembles, définition ensembliste de $\mathbf{N}$, nombres transfinis
- 1920-1930 : Heisenberg, Dirac : naissance de la mécanique quantique. Les nombres non-commutatifs envahissent la Physique
- Fin du XXe siècle : La géométrie non-commutative jette un nouveau pont entre l'algèbre et la géométrie.

<u>Références</u> :

- Bourbaki, Algèbre chapitres 1-3
- J. Dieudonné, Pour l'honneur de l'esprit humain, Hachette 1987
- G. Ifrah, Histoire universelle des chiffres, Robert Laffont 1994
- M. Ruhlen, L'origine des langues, sur les traces de la langue mère, Belin 1997
- B.L. Van der Waerden, A history of Algebra, Springer-Verlag 1985

<u>Référence électronique</u> :

http ://www-groups.dcs.st-and.ac.uk/˜history