DOMAINE : Non olympique
NIVEAU : Intermédiaire
CONTENU : Cours

AUTEUR : Louis NEBOUT
STAGE : Montpellier 2013

Dénombrabilité

Le but de cette séance d'ouverture est d'essayer de compter jusqu'à l'infini, en vous introduisant à la théorie des cardinaux, une des révolutions majeures de l'histoire des mathématiques, développée par Cantor autour de la fin du 19° siècle. Nous donnerons, quand c'est possible, les définitions précises de ces quantités infinies, puis nous essaierons de les comparer, et de dénombrer certains ensembles infinis usuels.

- Cardinaux : définitions -

On commence par un rappel :

Définition 1. Une application $f : E \longrightarrow F$ est dite
- *injective* si deux éléments distincts de E ont des images distinctes par F (i.e. f "sépare" les éléments de E)
- *surjective* si tout élément de F a un antécédent par f (i.e. f "recouvre" F tout entier)
- *bijective* si elle est injective et surjective (i.e. f "couple" les éléments de E avec ceux de F).

J'ai 7 poules. Que se passe-t'il quand je compte mes poules ? Je choisis une poule, je décrète que c'est la poule numéro 1, puis je prend une autre poule, qui sera la poule numéro 2, et ainsi de suite jusqu'à 7. Implicitement, j'ai donc construit une bijection entre mon ensemble de poules et l'ensemble $\{1, \dots, 7\}$. D'où la définition :

Définition 2. Soit n un entier non nul. On dit qu'un ensemble E a n éléments s'il est en bijection avec $\{1, \dots, n\}$. On dit alors que le cardinal de E est n, et on note cela $|E| = n$. (On dit qu'un ensemble est de cardinal 0 s'il est en bijection avec l'ensemble vide, c'est-à-dire s'il est vide).

L'avantage de ce point de vue est qu'il s'étend sans efforts aux ensembles infinis ! Ainsi :

Définition 3. Soient E et F deux ensembles (potentiellement infinis).
- Si E et F sont en bijection, on dit qu'ils ont même *cardinal*, on note cela $|E| = |F|$.
- S'il existe une injection de E dans F, on dit que le cardinal de E est plus petit que celui de F, on note cela $|E| \leqslant |F|$.
- S'il existe une surjection de E vers F, on dit que le cardinal de E est plus grand que celui de F, on note cela $|E| \geqslant |F|$.

La notation $\leqslant$ n'est pas innocente, et on voit d'ailleurs que cette relation vérifie plusieurs des propriétés naturelles d'une relation de comparaison, par exemple $|E| \leqslant |F|, |F| \leqslant |G| \Rightarrow |E| \leqslant |G|$ en composant les injections. Le théorème suivant regroupe plusieurs résultats montrant que notre relation vérifie toutes les propriétés naturelles d'une relation de comparaison, et on peut donc la manipuler comme un ordre classique.

Théorème 4. Soient E et F deux ensembles (potentiellement infinis).
- S'il existe une injection de E dans F, et une injection de F dans E, alors E et F sont en bijection, autrement dit : $|E| \leqslant |F|, |F| \leqslant |E| \Rightarrow |E| = |F|$.
- Il existe une injection de E dans F si et seulement s'il existe une surjection de F dans E, autrement dit : $|E| \leqslant |F| \Leftrightarrow |F| \geqslant |E|$.
- Il existe toujours soit une injection de E dans F, soit une injection de F dans E, autrement dit : $|E| \leqslant |F|$ ou $|F| \leqslant |E|$.

Démonstration. Je vais juste prouver la première propriété, connue sous le nom de théorème de Cantor-Bernstein. Cette preuve est un peu technique et peut tout à fait être sautée. On cherche à construire une bijection φ de E dans F. Soit f l'injection de E dans F et g celle de F dans E. Une idée est de partir de f. On décrète que φ vaut f sur E. Il existe alors des points de B non atteints par φ : les points y hors de $f(E)$, pour eux on a envie de se servir de g, par exemple dire que, $\varphi(g(y)) = y$. Le problème, c'est qu'après cette modification le point $f(g(y))$ n'est plus atteint par φ, il faudrait donc trouver un moyen de répéter à volonté le processus.

Une façon de faire cela proprement est d'introduire des « chaînes ». Soit x dans E, on définit la chaîne de x comme étant la suite (x_n) définie par récurrence par $x_{2k+1} = f(x_{2k})$ et $x_{2k} = g(x_{2k-1})$. Si cette suite revient sur x, on dit que la chaîne de x est une boucle. Si ce n'est pas une boucle, on l'étend le plus possible sur les entiers négatifs : si x a un antécédent y par g, il n'en a qu'un, et on

pose $x_{-1} = y$. Puis on cherche un antécédent de x_{-1} par f, et ainsi de suite. Si ce processus bloque au bout d'un moment (si on tombe sur un élément n'ayant pas d'antécédent), on dit que la chaîne a une origine, sinon on dit qu'elle est infinie.

On va définir φ directement sur les chaînes : si la chaîne de x est une boucle ou est infinie, on prend pour $\varphi(x)$ l'élément suivant x dans la chaîne (c'est-à-dire $f(x)$). Si la chaîne a une origine et que cette origine est dans E, on prend pour $\varphi(x)$ l'élément suivant x dans la chaîne, mais si l'origine de la chaîne est dans F, on prend pour $\varphi(x)$ l'élément précédent x dans la chaîne (si on ne faisait pas cela, l'origine de la chaîne n'aurait pas d'antécedent par φ). On vérifie qu'une telle définition fonctionne. □

Les deux propriétés qui restent sont plus méchantes car elles nécessitent de faire appel à l'axiome du choix.

- Petite digression : l'axiome du choix -

Les axiomes sont le point de départ de toute théorie mathématique : ce sont des propriétés basiques et naturelles que l'on admet et dont on se sert ensuite pour déduire toute la théorie. Un exemple célèbre est l'axiome des parallèles de la géométrie euclidienne : par tout point passe une unique parallèle à une droite donnée. Quand on fait de la théorie des ensembles, et que l'on dispose d'un ensemble non vide E, on a bien sûr le droit de se fixer un élément $x \in E$, puis de travailler dessus. L'axiome du choix affirme que l'on peut faire une infinité de tels choix simultanément :

Axiome 1. Soit $(E_i)_{i\in I}$ une famille d'ensembles. Alors il existe une fonction f, appelée fonction de choix, associant à chaque i un élément de E_i.

Cela paraît totalement évident, à tel point que les mathématiciens ont mis longtemps à se rendre compte que cette propriété était indépendante de tous les axiomes usuels de la théorie des ensembles, et méritait donc d'être érigée en axiome. Pourtant, quand on y réfléchit, la situation est loin d'être aussi claire.

Supposons que les E_i soient des parties de $\mathbb{N}$. Dans ce cas, nul besoin de faire appel à l'axiome du choix : on peut définir une fonction de choix explicite, par exemple la fonction "plus petit élément". Ainsi, dans ce cas, l'axiome du choix se déduit des axiomes de base.

La situation se complique quand on travaille avec des ensembles plus compliqués, comme par exemple celui des réels. Supposez que les E_i soient biscornus, et essayez de trouver une fonction de choix universelle, "canonique". Celle utilisée pour les entiers, "plus petit élément", ne marche plus, un E_i n'ayant pas forcément de plus petit élément ! On se convainc rapidement qu'un choix "canonique" est impossible. Obtenir une fonction de choix nécessite alors de faire une infinité de choix totalement arbitraires, une chose qu'il n'est pas possible de faire avec les autres axiomes. Faire appel à l'axiome du choix, en quelque sorte, c'est introduire une grande quantité d'arbitraire dans les mathématiques.

Du coup, l'axiome du choix a un statut un peu à part parmi les autres axiomes. Quand on l'utilise, on se retrouve avec des objets vraiment moches, non explicites, et donc non manipulables par un ordinateur. Un exemple d'objet ainsi construit est une solution non continue de l'équation de Cauchy $f(x+y) = f(x) + f(y)$. Les mathématiciens modernes utilisent régulièrement l'axiome du choix, notamment en analyse, ou comme ici en théorie des ensembles, mais essayent dès que possible de s'en passer, afin que les mathématiques restent accessibles d'un point de vue algorithmique.

- Ensembles dénombrables -

Un ensemble ayant un cardinal entier est appelé un ensemble fini. Cette section est consacrée à l'étude de ce qu'il se passe juste après les cardinaux entiers. On montre en effet que :

Définition 5. Le cardinal de $\mathbb{N}$ est le plus petit cardinal infini. Un ensemble ayant même cardinal que $\mathbb{N}$ (i.e. en bijection avec $\mathbb{N}$, i.e. dont on peut énumerer les éléments) est appelé un ensemble *dénombrable*.

Démonstration. Soit E un ensemble infini. Il s'agit de montrer que $\mathbb{N}$ s'injecte dans E. Soit x_1 un élément de E. Comme E est infini, $E\backslash\{x_1\}$ est non vide, soit x_2 dans $E\backslash\{x_1\}$. On repète le procédé : supposons $\{x_1, \dots, x_n\}$ construits, alors on prend pour x_{n+1} un élément de l'ensemble non vide $E\backslash\{x_1, \dots, x_n\}$. L'application $i \mapsto x_i$ est alors l'injection recherchée.

Il y a une subtilité. Dire "on repète le procédé" n'a pas vraiment de sens dans le cadre de la théorie des ensembles. En fait, une des formes de l'axiome du choix, appelée axiome des choix dépendants, dit précisément qu'une telle construction est possible. (On choisit une infinité de x_i à la fois, il est donc naturel de devoir faire appel à l'axiome du choix). □

Ainsi, beaucoup d'ensembles qui pourraient sembler à première vue plus petits que $\mathbb{N}$, par exemple l'ensemble des nombres pairs, celui des nombres premiers, ou celui des puissances de 2, sont en fait dénombrables, et donc de même taille que $\mathbb{N}$. Le phénomène inverse est vrai : beaucoup d'ensemble paraissant plus gros que $\mathbb{N}$ sont en fait dénombrables.

Théorème 6. Les ensembles suivant sont dénombrables : $\mathbb{Z}$, $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$, $\mathbb{Q}$.

Démonstration. Une énumération des éléments de $\mathbb{Z}$ est facile à trouver : prendre par exemple $(0, 1, -1, 2, -2, \ldots)$. On peut la traduire en une bijection explicite : $n \mapsto (-1)^{n+1}\lfloor (n+1)/2 \rfloor$.

Pour $\mathbb{N}^2$, on énumère les couples d'entiers diagonales par diagonales :

$$((0,0),(0,1),(1,0),(0,2),(1,1),(2,0),(0,3),(1,2),(2,1),(3,0),(0,4),\ldots).$$

On peut à nouveau expliciter la bijection : la diagonale de $\mathbb{N}^2$ des points (m, n) tels que $m + n = k$ contient $k + 1$ points. Ainsi, les diagonales en dessous de celle contenant le point (m, n) contiennent au total $1+2+\ldots+(m+n)$ points, et la bijection est donnée par $(m, n) \mapsto (m + n)(m + n + 1)/2 + m + 1$.

Pour $\mathbb{Q}$, il est plus difficile d'obtenir une bijection explicite. On peut par contre remarquer que $\mathbb{Q}$ s'injecte dans l'ensemble $\mathbb{N}^2$ (par l'application associant à un rationnel son numérateur et son dénomniateur, lorsque la fraction est sous forme réduite). Ainsi, le cardinal de $\mathbb{Q}$ est infini, mais plus petit que celui d'un ensemble dénombrable, et le résultat précédent dit que $\mathbb{Q}$ est dénombrable.

Pour ceux que cela intéresse, le livre "Raisonnements divins" donne des exemples de bijection explicite entre $\mathbb{Q}$ et $\mathbb{N}$. Par exemple, considérons la fonction

$$f \mapsto \frac{1}{\lfloor x \rfloor + 1 - \{x\}},$$

et définissons une suite par $x_0 = 1$, et $x_{n+1} = f(x_n)$. Alors $n \mapsto x_n$ est une bijection de $\mathbb{N}$ dans $\mathbb{Q}$. Une autre façon de le voir est de considérer l'arbre binaire dont la racine est étiquetée par le rationnel $1/1$ (écrit sous cette forme), et dont les deux fils d'un sommet étiqueté par a/b sont étiquetés par $a/(a+b)$ et $(a + b)/b$. Alors, quand on parcourt l'arbre de gauche à droite puis de bas en haut, on retrouve la suite précédente. Les preuves ne sont pas très difficiles à faire, et constituent un exercice amusant. □

L'argument utilisé pour prouver que $\mathbb{Q}$ est dénombrable permet en outre d'obtenir le résultat suivant, trè utile en pratique, et dont nous donnons immédiatement un exemple d'application :

Lemme 7. Toute union dénombrable d'ensembles dénombrables est dénombrable.

Démonstration. Soit $(U_i)_{i\in A}$ une famille dénombrable d'ensembles dénombrables, et soit f une bijection de A dans $\mathbb{N}$ et $(f_i)_{i\in\mathbb{N}}$ une famille de bijections de $\mathbb{N}$ dans U_i (pour construire cette famille, on a utilisé l'axiome du choix). Il s'agit de prouver que

$$U := \cup_{i\in A} U_i$$

est dénombrable. Or cet ensemble s'injecte dans $\mathbb{N}^2$ par l'application envoyant un élément x de U sur le couple $(f_i(x), f(i))$, où i est l'élément de A tel que x appartienne à U_i et que $f(i)$ soit minimal. □

Proposition 8. L'ensembles des nombres algébriques, c'est-à-dire des réels racines d'une équation polynomiale à coefficients dans $\mathbb{Q}$, est dénombrable.

Démonstration. L'ensemble des polynômes de degré d à coefficients dans $\mathbb{Q}$ est dénombrable (il est en bijection avec $\mathbb{Q}^{d+1}$), donc par le lemme l'ensemble des polynômes de degré quelconque à coefficients dans $\mathbb{Q}$ est lui aussi dénombrable. L'ensemble des racines d'un tel polynôme est fini, donc l'ensemble des nombres algébriques est une union dénombrable d'ensembles finis, et la preuve du lemme s'adapte sans difficultés pour prouver qu'un tel ensemble est dénombrable. □

Remarques 9. Le lemme admet une généralisation bien plus puissante : si E et F sont des ensembles tels que $|F| < |E|$ (et tel que F soit non vide), alors $|E^F| = |E|$, où E^F est l'ensemble des applications de F dans E.

- Au delà de l'infini -

La question qui vient ensuite est la suivante : y'a-t'il des ensembles plus gros que celui des entiers ? La réponse est oui, et pas besoin de chercher bien loin pour les trouver. Le résultat suivant de Cantor, qui ne ressemblait à aucun autre résultat connu à l'époque, a eu l'effet d'un coup de tonnerre dans la communauté mathématique.

Théorème 10. $|\mathbb{R}| > |\mathbb{N}|$.

Démonstration. Nous allons montrer que $\mathbb{N}$ ne peut pas se surjecter dans $[0,1[$, et on aura ainsi $|\mathbb{N}| < |[0,1[| \leqslant |\mathbb{R}|$. Soit donc f allant de $\mathbb{N}$ dans $[0,1[$. Rappelons que tout réel de $[0,1[$ admet un unique développement décimal propre, c'est-à-dire un développement ne se terminant pas par une infinité de 9. On écrit les

développements décimaux propres de tous les $f(i)$, et on cherche à exhiber un réel n'apparaissant pas dans cette liste. Une façon naturelle de le faire est de choisir un réel ayant un développement décimal propre différent de tous ceux des $f(i)$. Une telle construction est donnée par ce que l'on appelle un argument diagonal :

Soit x_i un entier de $\{0, \dots, 8\}$, différent de la i-ième décimale du développement décimal propre de $f(i)$. Alors $0, x_1x_2 \dots$ est le développement propre d'un réel x, et x est différent de chacun des $f(i)$, puisque les i-ème décimales de leurs uniques développements propres sont différentes. Ainsi, x n'est pas dans l'image de f, et f n'est pas surjective. □

L'ensemble des réels est donc strictement plus gros que celui des entiers. On se demande donc s'il existe un ensemble de taille intermédiaire entre ces deux ensembles. Et bien, étrangement, cette propriété, connue sous le nom d'"hypothèse du continu", est indépendante de tous les axiomes de la théorie des ensembles, et il est donc impossible de répondre à cette question. Contrairement à l'axiome du choix, la véracité ou non de l'hypothèse du continu n'a que peu d'influence sur le reste des mathématiques, et il n'y a donc pas besoin d'introduire un nouvel axiome.

Voici un autre résultat de même nature que le précédent :

Proposition 11. Soit E un ensemble. On note $\mathcal{P}(E)$ l'ensemble des parties de E. Alors $|\mathcal{P}(E)| > |E|$. (remarque : $\mathcal{P}(E)$ est en bijection avec 2^E.)

Démonstration. Tout d'abord, E s'injecte dans $\mathcal{P}(E)$ par l'application associant à l'élément x le singleton $\{x\}$. Ainsi, $|\mathcal{P}(E)| \geqslant |E|$. Supposons maintenant l'existence d'une surjection f de E dans $\mathcal{P}(E)$. Définissons

$$U := \{x \in E | x \notin f(x)\}.$$

Comme f est surjective, on peut trouver u tel que $f(u) = U$. Cet élément u est-t'il dans U ? S'il l'était, u ne serait pas dans $f(u)$, donc dans U, ce qui est absurde, mais s'il ne l'était pas, alors cela voudrait dire que u est dans $f(u)$, c'est-à-dire U. Nous arrivons donc à un paradoxe, cela contredit l'existence de la surjection. □

- Conséquences de la non dénombrabilité de $\mathbb{R}$ -

Une conséquence importante est qu'une partie dénombrable de $\mathbb{R}$ ne peut être égale à $\mathbb{R}$ tout entier. Ainsi, il existe des réels qui ne sont pas algébriques !

On appelle un tel réel un nombre transcendant. L'existence de ces réels n'est pas une question facile : il est déjà relativement difficile d'expliciter de tels réels, il est encore plus difficile de prouver qu'un réel donné est transcendant. Encore aujourd'hui, on ne sait le faire que pour peu de réels. Et pourtant, le résultat de Cantor nous dit en deux lignes qu'il existe des nombres transcendants, et même pire que cela, qu'il en existe une infinité non dénombrable ! On peut aller plus loin.

En théorie des probabilités, on définit une quantité appelée la mesure de Lebesgue, notée λ, de telle sorte que, si on tire un réel au hasard dans $]0,1[$, ce réel est dans l'ensemble A avec probabilité $\lambda(A)$. Cette mesure vérifie les propriétés que l'on espére, par exemple, si $0 \leqslant a \leqslant b \leqslant 1$, alors $\lambda(]a,b[) = b - a$, λ est croissante pour l'inclusion, et la mesure de l'union disjointe de A et B est égale à la somme des mesures de A et de B.

On vérifie alors qu'un sous-ensemble E dénombrable de $]0,1[$ est toujours de mesure nulle ! Admettons qu'il soit mesurable, soit ϵ un réel strictement positif, et soit $(e_1, e_2, \ldots)$ une énumération des éléments de E. Pour tout i, on inclut e_i dans un intervalle ouvert de $]0,1[$ de mesure $\frac{\epsilon}{2^i}$. Alors, E sera inclus dans une réunion d'intervalles ouverts, de mesure totale inférieure à

$$\frac{\epsilon}{2^1} + \frac{\epsilon}{2^2} + \frac{\epsilon}{2^3} + \ldots = \epsilon.$$

Ainsi, la mesure de E est inférieure à ϵ, et ce pour tout ϵ, donc E est de mesure nulle.

Ainsi, un réel choisi au hasard dans $]0,1[$ sera transcendant avec probabilité 1 ! On pourrait montrer, comme on l'a fait avec les nombres algébriques, que l'ensemble des réels constructibles à la régle et au compas, ou que l'ensemble des réels définissables par une phrase (ou, si vous préférez , l'ensemble des réels calculables par un ordinateur) sont eux aussi dénombrables. Ainsi, un réel choisi au hasard sera avec probabilité 1 transcendant, non constructible et non calculable ! (Un bémol toutefois : comment chosir un réel au hasard ?) Cantor a ainsi réussi à prouver, par sa théorie des cardinaux, que la majorité des nombres réels seraient à tout jamais hors de portée des mathématiciens, et, ce faisant, a ébralé de manière irréversible la conception idéaliste qu'avaient alors les gens envers la portée des mathématiques, une révolution qui serait achevée des années plus tard par Gödel, mais c'est une autre histoire…