L'espace mathématique comme un réseau

Variation sur une conférence de Poincaré, un siècle plus tard…

L'Imagination et l'Intuition dans les Sciences
Phantasie und Intuition in den Wissenschaften

Institut de France
Fondation Simone et Cino del Duca
Paris, 7 et 8 novembre 2008

Étienne Ghys

Il y a cent ans et quelques mois, Henri Poincaré donnait à Paris une conférence restée célèbre sur *l'invention mathématique*[1] dans laquelle il décrivait les circonstances de sa découverte des *groupes fuchsiens*[2]. Il mettait ainsi en évidence le rôle conjoint de la logique *consciente* et de l'intuition *inconsciente*.

« C'est par la logique que l'on prouve et par l'intuition que l'on découvre ».

Mon but, modeste, est de revisiter cette conférence de Poincaré, à la lumière du fonctionnement d'*internet*. Il n'y aura aucune nouveauté, sauf peut-être vers la fin, et je vous demande de l'indulgence pour un mathématicien qui se risque en dehors de ses terrains de chasse habituels.

On compare souvent les mathématiques à un jeu d'échecs. Une position, l'hypothèse, est donnée. Les règles du jeu sont connues et imposées ; et il s'agit de manœuvrer les pièces pour aboutir à une conclusion : échec et mat ! C'est bien sûr une analogie très réductrice mais elle n'est pas sans valeur. Bien souvent d'ailleurs notre système éducatif limite les compétences mathématiques qu'on enseigne à ce

---

[1] HENRI POINCARÉ : *L'invention mathématique*, conférence à l'Institut de Psychologie, 23 mai 1908, *L'Enseignement mathématique,* volume 10, pp.357-371 (1908). Disponible sur http://retro.seals.ch

[2] *« Un soir, je pris du café noir contrairement à mon habitude ; je ne pus m'endormir ; les idées surgissaient en foule ; je les sentais comme se heurter, jusqu'à ce que deux d'entre elles s'accrochassent pour ainsi dire pour former une combinaison stable. Le matin j'avais établi l'existence d'une classe de fonctions fuchsiennes, celles etc. »*

genre de jeu-exercice. Voyez par exemple l'une des choses que vous deviez savoir faire si vous vouliez entrer à l'École Normale Supérieure l'année dernière. On donne une hypothèse, une conclusion, et c'est à vous de trouver le chemin. Bien sûr, la difficulté c'est qu'il y a en général une grande quantité de chemins possibles et que le plupart ne mènent à rien, ou ne nous mènent pas là où nous voulons aller. Mais lisons plutôt ce qu'écrit Poincaré.

Qu'est-ce, en effet, que l'invention mathématique ? Elle ne consiste pas à faire de nouvelles combinaisons avec des êtres mathématiques déjà connus. Cela, n'importe qui pourrait le faire ; mais les combinaisons que l'on pourrait faire ainsi seraient en nombre fini, et le plus grand nombre est absolument dépourvu d'intérêt. Inventer, cela consiste précisément à ne pas construire les combinaisons inutiles et à construire celles qui sont utiles et qui ne sont qu'une infime minorité. Inventer, c'est discerner, c'est choisir.

[L'inventeur] fait penser à un acheteur à qui l'on présente un grand nombre d'échantillons, qui les examine l'un après l'autre de façon à faire son choix. Ici les échantillons seraient tellement nombreux qu'une vie entière ne suffirait pas pour les examiner. Ce n'est pas ainsi que les choses se passent. Les combinaisons stériles ne se présenteront même pas à l'esprit de l'inventeur. Dans le champ de sa conscience n'apparaîtront jamais que les combinaisons réellement utiles, et quelques autres qu'il rejettera, mais qui participent un peu des caractères de combinaisons utiles.

Tout se passe comme si l'inventeur était un examinateur du deuxième degré, qui n'aurait plus à interroger que les candidats déclarés admissibles après une première épreuve.

Ainsi donc, il faut que notre inconscient fasse un tri. Il existe une situation analogue en informatique. Lorsque j'ouvre http://maps.google.com/ et que je cherche l'itinéraire entre le siège de *l'Académie Leopoldina* et *l'Académie des sciences* à Paris, le logiciel m'affiche immédiatement le meilleur chemin. Lui aussi fait des choix. Il en fait beaucoup, mais comment fait-il pour sélectionner en si peu de temps le meilleur chemin ?

Permettez-moi de vous présenter en quelques mots l'algorithme utilisé, qui porte le nom peu poétique de *A étoile*[3]. La première idée, mauvaise, consisterait à explorer systématiquement tous les chemins qui partent de l'Académie Leopoldina jusqu'au moment où on aboutit quai de Conti. On le conçoit, cette stratégie fonctionne mais elle est passablement aveugle et prendrait un temps de calcul peu raisonnable ! Il paraît immédiatement évident à notre intuition qu'il est préférable par exemple de partir vers l'ouest ! L'algorithme *A étoile*, que je n'ai pas le temps de décrire en détail, n'explore *pas* tous les chemins. Après avoir examiné progressivement un certain nombre de chemins partiels partant de l'Académie Leopoldina, on explore toutes les manières possibles de les prolonger disons jusqu'à la prochaine ville x et, pour chaque tentative de prolongement jusque x, on évalue deux nombres. Le premier $T(x)$ est le temps nécessaire pour aller de Leopoldina jusqu'en x, sur le chemin que en cours d'exploration. Et le second $H(x)$ est ce que les informaticiens appellent *l'heuristique* de x. C'est une estimation du temps nécessaire pour aller de x jusqu'au quai de Conti. On ne connaît pas le plus court chemin de x à Paris, sinon le problème serait résolu, mais on peut en avoir une estimation. Par exemple, on peut estimer ce temps en calculant la distance *à vol d'oiseau* entre x et Paris, qui nous donnera une indication grossière. L'algorithme *A étoile* choisira alors de continuer son chemin en allant à la ville x qui minimise la somme $T(x)+H(x)$. En quelque sorte, l'algorithme aveugle qui explorerait tous les chemins sans exception est guidé par l'heuristique qui lui souffle qu'il n'est peut-être pas une bonne idée de partir vers l'est ! Dans la pratique, cet algorithme va très vite d'un point à un autre et n'explore qu'un nombre de possibilités extrêmement petit alors que l'algorithme aveugle explore tout systématiquement.

On voit que la situation est proche du processus décrit par Poincaré dans lequel l'intuition inconsciente choisit pour le chercheur un petit nombre de chemins qu'il peut alors essayer. On voit même que le mot choisi par les informaticiens « heuristique » n'est pas si loin de l'intuition.

---

[3] Voir par exemple http://en.wikipedia.org/wiki/A*_search_algorithm

Un problème important se présente cependant. Si l'on veut choisir le bon chemin dans l'espace mathématique, quelle serait la fonction heuristique *H(x)* qui estime le temps nécessaire pour aboutir à un énoncé mathématique *x* en partant d'une situation donnée ? Y a-t-il quelque chose comme la distance à vol d'oiseau entre deux théorèmes[4] ? Et si oui, comment diable notre inconscient peut-il l'évaluer ? Mais plus sérieux encore : il est bien rare qu'un mathématicien travaille comme le jour du concours d'entrée à l'École Normale Supérieure : il sait (en général) où il est, il connaît les hypothèses pourrait-on dire, mais il ne sait pas exactement où il veut aller. Ce qu'il veut, c'est visiter des contrées intéressantes et *mathématiquement pertinentes*. Il s'agit plus de l'explorateur qui visite un continent inconnu que du touriste équipé de son guide du routard ! Qu'est-ce qui donne l'intuition à l'explorateur que cette nouvelle vallée qui se présente à lui est prometteuse ?

Poincaré propose un critère de sélection par lequel l'inconscient choisit : selon lui le choix est esthétique. Voyons ce qu'il écrit :

> De sorte que nous arrivons à la conclusion suivante : Les combinaisons utiles, ce sont précisément les plus belles, je veux dire celles qui peuvent le mieux charmer cette sensibilité spéciale que tous les mathématiciens connaissent, mais que les profanes ignorent au point qu'ils sont souvent tentés d'en sourire.

Même s'il m'en coûte de critiquer Poincaré, il me semble qu'il y a là une « tricherie » puisque Poincaré ne donne absolument aucune définition de ces « belles combinaisons » que « tous les mathématiciens connaissent ». Oui, il me semble que Poincaré évacue le problème en utilisant un mot qu'il ne définit pas : un péché mortel pour un mathématicien !

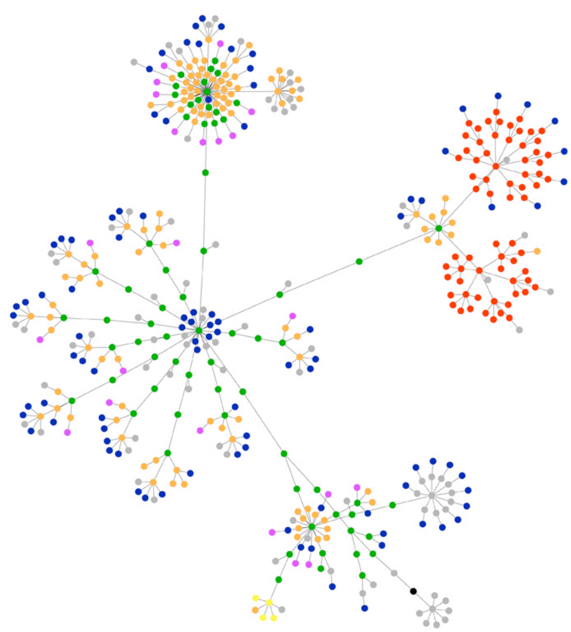
Je vous propose une deuxième analogie informatique. Comparons l'espace mathématique dans lequel le chercheur évolue et le réseau de la *Toile*, du *Web* si

---

[4] Ah ! que les mathématiciens rêveraient de disposer d'une telle fonction leur disant à l'avance combien de temps ils doivent travailler avant d'aboutir à un théorème convoité !

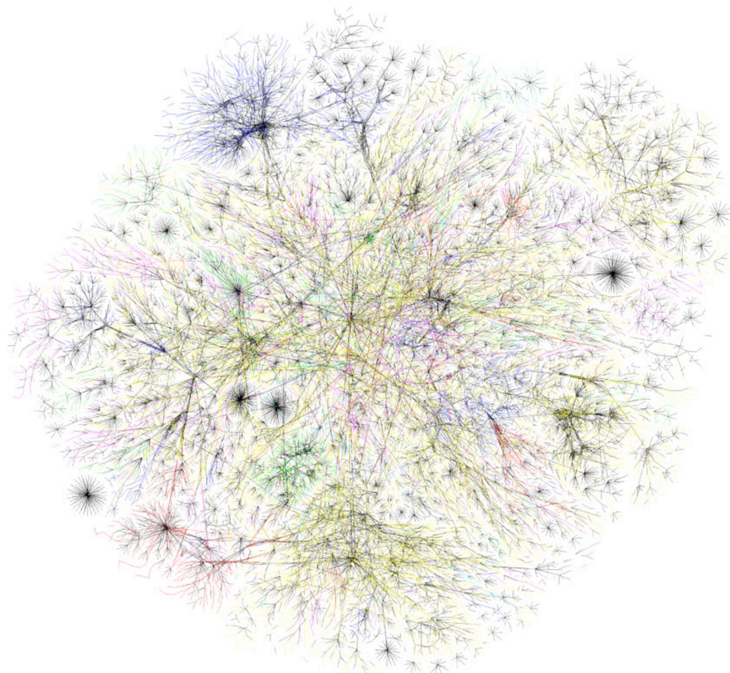
vous préférez[5]. La Toile est un immense réseau possédant un grand nombre de nœuds, un pour chaque page internet dans le monde (on dit qu'il y en a environ cent milliards aujourd'hui, du même ordre de grandeur que le nombre de neurones dans un cerveau humain) et dans lequel une connexion lie deux nœuds si l'on passe de l'un à l'autre en cliquant une fois sur sa souris. En moyenne une page est connectée à cinq ou six autres (beaucoup moins que les dizaines de milliers de neurones connectés à un seul neurone). On peut imaginer que le monde mathématique est un peu comme cela : une gigantesque collection de faits mathématiques interconnectés de manière complexe par des processus élémentaires, des syllogismes pour simplifier. Mais permettez-moi d'exprimer ici ce que presque tous les mathématiciens ressentent : l'espace mathématique est incommensurablement plus vaste que la Toile, et même que notre cerveau. Alors, comment naviguer dans un tel espace ? Eh bien, en surfant comme sur internet, et en utilisant *Google* !

Voici d'abord quelques images, simplistes, qui peuvent vous donner une idée de la géométrie de la Toile. Ici, j'utilise un logiciel qui montre la structure dans le voisinage de ma propre page personnelle. Ma page est au centre, et on voit mes voisins et les liens qui nous unissent. Vous remarquez cette structure étonnante, pleine de bourgeonnements.

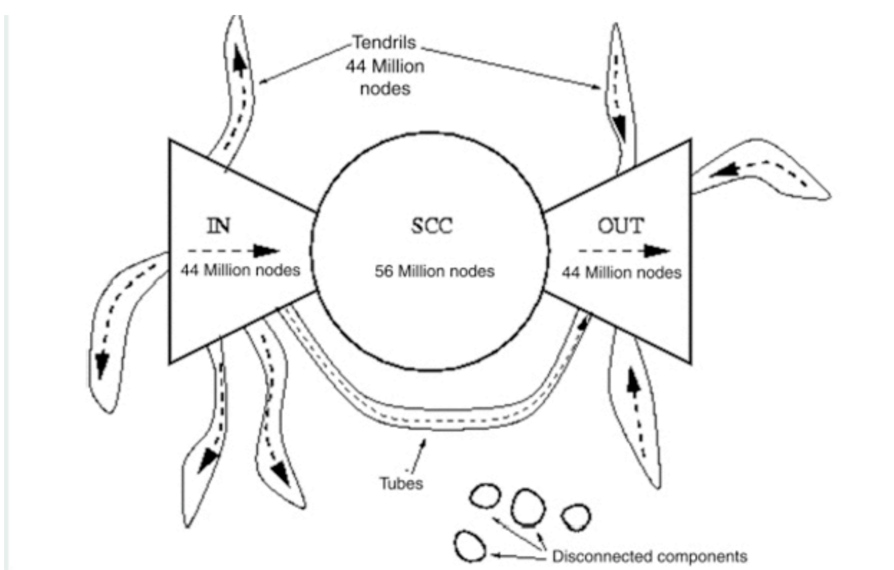


[5] Voir par exemple : ANTHONY BONATO, *A course on the Web graph*, Graduate Studies in Mathematics, volume 89, American Mathematical Society (2008).

Voici une image globale de la Toile, dont je ne sais pas trop ce qu'elle signifie à vrai dire. C'est peut-être une vision d'artiste qui illustre encore ces bourgeonnements et qui rappelle un peu les feux d'artifice du 14 juillet !



Et enfin, voici un schéma intéressant déjà ancien, dit du "nœud papillon" qui donne une structure globale.



En gros, les pages internet sont de quatre sortes. Il y a le SCC, le « strongly connected core » qui est vraiment le cœur actif, où tout communiquent avec tout. Il y a la partie OUT, formée des pages que tout le monde cite, mais qui ne citent personne. Il y a la partie IN, formée des pages qui citent les autres mais que personne ne cite. Et puis, il y a des dendrites, des tubes et même des îlots isolés du reste du cyber-espace. Il me semble que si je devais tracer une carte du monde mathématique, je dessinerais exactement la même chose. Certains domaines intéressent tout le monde mais ne s'intéressent à personne, d'autres au contraire n'intéressent personne, d'autres sont au cœur des maths, et enfin nous avons aussi nos dendrites, nos tubes et même nos îles isolées du reste du monde.

Comment se repérer dans une telle immensité ? Quelle heuristique utiliser dans ce réseau gigantesque ? Quelles sont les pages qui sont pertinentes ? L'idée géniale de *Google* est d'utiliser une méthode que les mathématiciens connaissent bien : la distribution qu'ils appellent stationnaire, ou encore harmonique ; et que *Google* appelle le *PageRank* qui attribue une « note » à chaque page internet mesurant son « intérêt »[6]. L'idée est simple et d'ailleurs *Google* lui-même la qualifie de « démocratique »[7] ! La note est recalculée, par exemple tous les jours, de la manière suivante. Au début, toutes les pages ont la note 1 par exemple. Puis, le lendemain, on examine pour chaque page combien d'autres pages pointent vers elle. Plus il y en a et plus on peut penser que la page considérée intéresse les internautes. On attribue alors à la page la somme des notes des pages qui la citent : par exemple 9 si 9 pages la citent. Mais le jour suivant, les notes des diverses pages ne sont plus toutes les mêmes si bien que si une page est citée par exemple par trois autres pages dont les notes sont 1, 5 et 50, on lui attribuera donc la nouvelle note 56. Drôle de démocratie que celle de *Google* : les gens considérés comme importants ont des voix qui comptent plus que les autres.... Par ce processus itératif, chaque page a donc un *PageRank* qui évolue de jour en jour d'ailleurs puisque la Toile n'est pas un réseau statique. Voilà l'heuristique qui guide le surfer sur internet.

Les ordinateurs énormes de *Google* tournent jour et nuit et visitent le monde entier, sans relâche, aveuglément, inconsciemment oserais-je dire, en tâche de fond comme on dit en informatique. Lorsque je cherche des informations sur l'académie des sciences par exemple, cette recherche est consciente et je tape quelques mots clés dans *Google*. Alors, *Google* utilise son travail inconscient, il connaît les pages intéressantes puisqu'il a calculé inconsciemment les *PageRank* et voilà : il a choisi

[6] Voir par exemple http://www.ams.org/featurecolumn/archive/pagerank.html
[7] Selon *Google*, « PageRank relies on the uniquely democratic nature of the Web by using its vast link structure as an indicator of an individual page's value ».

pour moi un petit nombre de pistes de recherche parmi l'immensité des possibles. Maintenant, mon conscient peut à nouveau travailler sur ce petit nombre de pistes. *Google* serait donc mon inconscient !

Puis-je reprendre la description par Poincaré du processus mental dans l'invention mathématique, en la reformulant dans notre nouvelle terminologie. Mon inconscient calcule en permanence, en tâche de fond, un *PageRank* pour tous les énoncés mathématiques qu'il visite de manière systématique en effectuant en quelque sorte un scrutin électoral à la *Google*. Mon conscient « tape » quelques mots clé en exprimant un souhait mathématique et voilà, mon inconscient me présente un petit nombre de pistes, que je dois ensuite examiner en détail. Je propose donc de remplacer les « belles combinaisons » que « tous les mathématiciens connaissent », chères à Poincaré, par un *PageRank* bien plus prosaïque, qui reste quand même bien mystérieux, et hypothétique, il faut bien le dire…

Cette présentation est-elle psychologiquement raisonnable ? Je n'en sais rien à vrai dire, puisque je n'ai aucune compétence dans ce domaine, mais elle me semble satisfaisante, tout au moins en première approximation. Si l'on quitte le niveau individuel d'un mathématicien et qu'on passe au niveau de la communauté des mathématiciens, il me semble en tous les cas que l'idée d'un *PageRank* est efficace pour comprendre l'opinion que nous nous faisons de telle ou telle partie des mathématiques. Le *Citation index* est bien sûr un outil primitif puisqu'il compte indistinctement toutes les citations, indépendamment de leur intérêt propre... Mais je pense que l'opinion de la communauté mathématique sur les diverses parties des maths est le résultat d'une réactualisation permanente et collective que nous faisons sur la base d'une « démocratie à la *Google* », avec ses électeurs plus puissants que d'autres ! Mais là encore le processus électoral n'est pas souvent explicite.

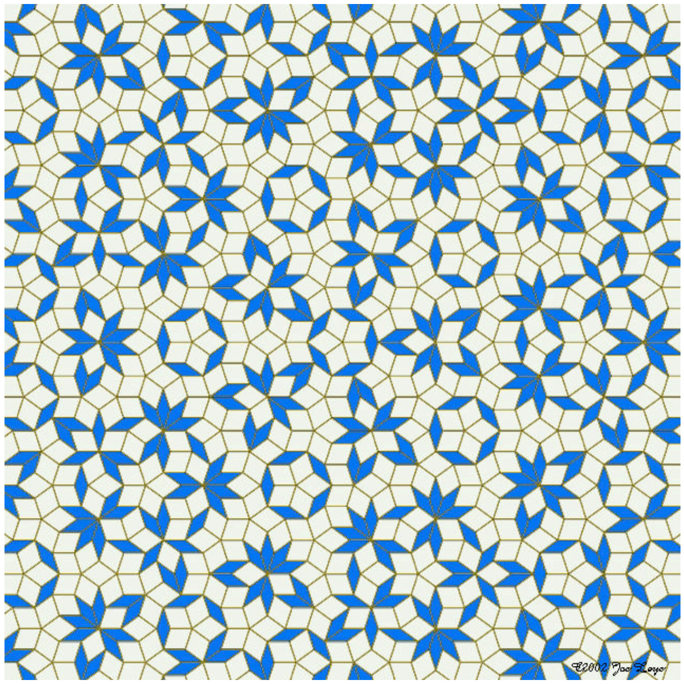
Évidemment, il ne faut pas limiter l'invention mathématique à ce travail inconscient. Poincaré nous met en garde : il faut travailler avant **et** après, et il ne faudrait pas oublier ces aspects de l'invention : le travail, la chance, le génie…

Il y a une autre remarque à faire au sujet des conditions de ce travail inconscient ; c'est qu'il n'est possible et, en tout cas, qu'il n'est fécond que s'il est, d'une part, précédé, et, d'autre part, suivi d'une période de travail conscient. Jamais (et les exemples que je vous ai cités le prouvent déjà suffisamment) ces inspirations subites ne se produisent sinon après quelques jours d'efforts volontaires, qui ont paru absolument infructueux et où l'on a cru ne rien faire de bon, où il semble qu'on a fait totalement fausse route. Ces efforts n'ont donc pas été aussi stériles qu'on le pense ; ils ont mis en branle la machine inconsciente, et sans eux elle n'aurait pas marché et elle n'aurait rien produit.

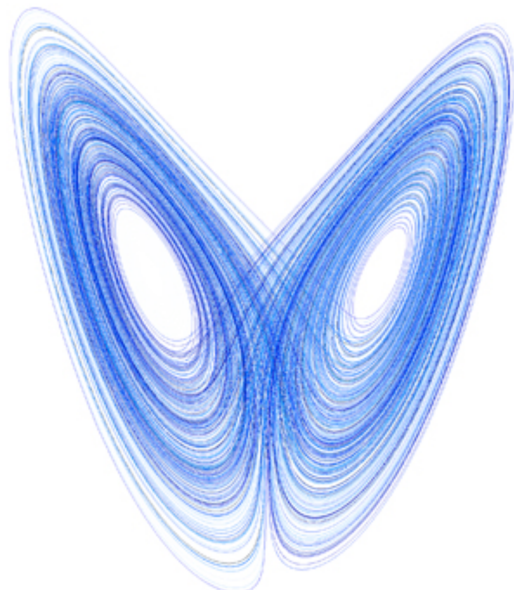
La nécessité de la seconde période de travail conscient, après l'inspiration, se comprend mieux encore. Il faut mettre en œuvre les résultats de cette inspiration, en déduire les conséquences immédiates, les ordonner, rédiger les démonstrations. Mais surtout il faut les vérifier.

Pour reprendre notre analogie avec la recherche sur *Google*, on pourrait dire qu'on n'aboutira pas à grand chose si on ne sélectionne pas avec le plus grand soin les mots clés employés, ou si l'on n'étudie pas sérieusement les pistes proposées par *Google.*

Je voudrais terminer par une élucubration (de plus !). La description précédente a deux inconvénients sérieux. D'abord, elle ne donne pas de rôle à *l'analogie* qui est pourtant reconnue par tous comme l'un des moteurs de l'intuition. D'autre part, les ordinateurs de *Google* visitent en fait *presque toutes* les pages dans le monde, car finalement le cyberespace n'est pas si grand... Par contre, je vous l'ai dit, le monde mathématique est bien plus vaste, et pour l'essentiel inexploré, si bien que mon inconscient n'a pas pu évaluer l'intérêt d'endroits qu'il n'a pas encore visité. L'explorateur-chercheur qui découvre une vallée inconnue lui rappelant une autre vallée qu'il a déjà explorée et qui l'avait mené à de vastes prairies, aura évidemment envie d'explorer plus avant cette nouvelle vallée. Il ne base pas cela sur un quelconque *PageRank* puisqu'il s'agit de sites qui n'ont pas encore été visités. Ma dernière élucubration consiste alors à postuler que le monde mathématique est *quasi-périodique*, un peu comme un de ces célèbres *pavages de Penrose* qui n'ont pas de symétrie véritable mais dans lesquels on retrouve sans cesse des sous-structures analogues, quasi-périodiquement.



Les mathématiciens le savent bien, la quasi-périodicité est une incarnation de la *compacité* au sens mathématique du terme. Il me semble alors que, plutôt que de considérer l'espace mathématique comme un réseau très grand, gigantesque, voire infini, je préfère l'imaginer comme un espace compact sur lequel un certain nombre d'opérations élémentaires agissent — disons pour simplifier les syllogismes — et dont les trajectoires décrivent le cheminement du mathématicien. L'image que je me ferais de l'espace mathématique serait alors plus proche de cette figure classique d'un attracteur dans lequel les trajectoires tournent de manière quasi-périodique. Une trajectoire infinie, confinée dans un espace compact est contrainte à passer et repasser sans cesse dans des situations presque identiques, des situations analogues pourrait-on dire, tout en étant différentes : c'est ce qu'on appelle la *récurrence de Poincaré.*



On entre alors dans le domaine de la théorie des *systèmes dynamiques* et de la *théorie ergodique*, dans laquelle les mesures stationnaires jouent un rôle central, si bien que l'idée d'un *Google* démocratique subsiste mais, cette fois, on peut tirer profit des *analogies* matérialisées par les récurrences, à la Poincaré. Mais je m'égare, et même si j'entre dans mon propre domaine de compétence mathématique, ce type d'interprétation psychologique est bien sûr de l'ordre de l'élucubration.

Étienne Ghys
CNRS-ENS Lyon
46 Allée d'Italie
69364 Lyon Cedex

etienne.ghys@umpa.ens-lyon.fr