

naïvement, même en mettant tous les dispositifs de calcul terrestres à la tâche, n'aurait même pas permis de résoudre le problème. Réduire le nombre de grilles à étudier est le premier point où les mathématiques ont aidé l'équipe de G. McGuire.

Si, dans une grille complète de sudoku, on change tous les 1 en 9 et tous les 9 en 1, on obtient une autre grille complète de sudoku. Si la première admet des problèmes corrects à 16 données, c'est aussi le cas de la seconde : partant d'un problème correct pour la première, on change les 1 en 9 et les 9 en 1, et on obtient un problème correct pour la seconde. Il suffit donc d'étudier l'une des deux. L'échange des 1 et des 9 n'est pas la seule transformation qui conduit à trouver des grilles équivalentes. En fait, toute permutation des neuf chiffres 1, 2, ..., 9 permet, à partir d'une grille complète, d'en construire une autre équivalente pour notre problème. Pour chaque grille, ce procédé donne $9! = 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 = 362\,880$ grilles équivalentes. Voilà déjà une belle réduction du calcul à entreprendre.

Cependant, la permutation des chiffres n'est pas la seule façon d'obtenir des grilles équivalentes. Voici la liste de quelques autres opérations qui, appliquées à une grille complète, donnent des grilles équivalentes.

- Choisir deux bandes horizontales et les échanger. Une bande horizontale est une série de trois carrés 3×3 placés les uns à côtés des autres (il y en a trois : une en haut, une au milieu, une en bas).
- Échanger deux bandes verticales. Une bande verticale est une série de trois carrés 3×3 placés les uns au-dessus des autres (il y en a trois : une à droite, une au centre, une à gauche).
- Échanger deux colonnes d'une même bande verticale.
- Échanger deux lignes d'une même bande horizontale.
- Transposer la grille (la ligne 1 devient la colonne 1, la ligne 2 devient la colonne 2, etc.).

Finalement, une grille est équivalente à un très grand nombre d'autres. Les grilles se regroupent en classes et il suffit de traiter une grille de chacune des classes.

9	3	7	8	5	6	2	4	1
5	6	2	1	9	4	3	8	7
4	8	1	2	7	3	5	6	9
8	2	3	6	4	7	9	1	5
6	1	5	9	3	2	4	7	8
7	4	9	5	8	1	6	2	3
3	7	8	4	6	9	1	5	2
1	9	6	7	2	5	8	3	4
2	5	4	3	1	8	7	9	6

DEUX « ENSEMBLES INÉVITABLES » sont indiqués (en bleu et en orange) dans cette grille. Chacun nécessite une donnée au moins pour que la solution au sudoku soit unique.

L'utilisation d'outils de la théorie des groupes, dont le lemme de Burnside (voir *Le miraculeux « lemme de Burnside » dans la rubrique de décembre 2006*), a permis de connaître le nombre exact de grilles complètes non équivalentes : il y en a $5\,472\,730\,538 \approx 5,5 \times 10^9$.

Glenn Fowler, chercheur aux Laboratoires AT&T, a constitué un catalogue complet des grilles en ne retenant qu'une seule grille par classe d'équivalence. Un algorithme de compression spécialement conçu pour les grilles de sudoku a permis de stocker tout le catalogue de G. Fowler sur six gigaoctets, ce qui est assez remarquable puisque cet algorithme permet de n'utiliser qu'un octet (8 bits) environ par grille.

Chercher les « ensembles inévitables »

Si l'on s'y prend sans finesse, pour chacune des grilles retenues, il faut envisager toutes les façons d'extraire 16 données. Or il y a $33\,594\,090\,947\,249\,085 \approx 3,4 \times 10^{16}$ façons de choisir 16 cases parmi 81. Cela conduirait à résoudre un grand nombre de problèmes de sudoku pour découvrir à chaque fois qu'il y a deux solutions ou plus. Ce nombre serait énorme, car égal à : $5\,472\,730\,538 \times 33\,594\,090\,947\,249\,085 = 183\,851\,407\,423\,359\,414\,572\,057\,730 \approx 1,8 \times 10^{26}$.

Nous savons que ce nombre est bien au-delà de ce qui est envisageable, même en utilisant toute la puissance de calcul disponible sur la planète. Les mathématiques à nouveau ont dû proposer une idée intéressante qui s'applique à d'autres problèmes, la « méthode des ensembles inévitables ». Expliquons cette idée sur l'exemple de la grille complète G ci-contre.

Pour construire un problème correct ayant pour solution cette grille G, il faut que l'une des données du problème soit l'une des quatre cases en bleu, en haut de la grille. En effet, il existe une autre grille complète G' très proche de G obtenue en échangeant seulement, dans les quatre cases en bleu, les 9 et les 5. Si les données de la grille partielle n'indiquent rien concernant ces quatre cases, il sera impossible

L'histoire de la conjecture

L'étude du problème du sudoku minimal par Gary McGuire et son équipe a commencé en 2005, par la mise au point d'un premier programme recherchant, pour une grille complète fixée, toutes les sous-grilles partielles correctes.

Cependant, c'est seulement en 2009, voyant que la solution du problème était peut-être atteignable, que G. McGuire et ses collègues ont entrepris des travaux systématiques. En effet, Gordon Royle avait à cette époque réuni une collection importante de problèmes corrects différents à 17 données. Il recherchait en même temps, mais en vain, un problème ne demandant que 16 données. Cette situation rendait plausible la conjecture que 16 données ne suffisent jamais. Une étonnante grille ayant 29 sous-grilles correctes à 17 données attira l'attention (ci-dessous). On ne connaît aujourd'hui encore aucune grille complète ayant 30 sous-grilles correctes à 17 données. Encouragé par ce grand nombre de sous-grilles correctes à

17 données, on espéra un moment en extraire une sous-grille correcte à 16 données. En vain, ce qui fut considéré comme un indice de plus en faveur de la conjecture.

Avant d'en arriver au résultat que 16 données sont insuffisantes, un calcul mené par une équipe de l'Université de Graz en Autriche montra, en 2008, que 11 ne suffisent pas. Le calcul entrepris par cette équipe pour montrer que 12 ne suffisent pas fut abandonné sans aboutir. On était donc loin du but. Rappelons toutefois que le raisonnement mathématique pur ne réussit pas mieux que 7 (voir l'encadré page 77) ; la démonstration par ordinateur était donc déjà en avance.

En 2008, une jeune fille de 17 ans proposa une preuve mathématique que 16 données sont insuffisantes, mais on découvrit

rapidement une erreur dans son raisonnement, que personne n'a su corriger.

En 2010, une équipe de l'Université Chiao Tung à Taiwan entreprit de s'attacher au problème des 16 données en utilisant une méthode de calcul distribué. Elle évalua alors qu'avec les outils dont elle disposait, le problème serait résolu en menant un calcul de 2 400 années-processeurs, par exemple, en faisant calculer 2 400 ordinateurs de table chacun pendant un an. Le calcul a abouti en septembre 2013 (après celui de G. McGuire). C'est la seule vérification à ce jour du théorème des 16 données.

Le calcul de G. McGuire et de son équipe se déroula sur toute l'année 2011, opérant l'équivalent de 800 années-processeurs de calculs. Pour chaque grille de sudoku complète, le temps de traitement a été enregistré. Il a fallu en moyenne 3,6 secondes par grille. L'article rendant compte du résultat a été publié en 2014.

de savoir si la solution est G ou G', et donc le problème sera incorrect. Il en va de même avec les six cases en orangé : permuter les deux rectangles de trois données {7, 2, 1} et {1, 7, 2} donne une grille G' qui n'est exclue que si l'on spécifie au moins une des six cases.

De tels ensembles se dénomment des *ensembles inévitables*. Si, pour une grille complète G, on en trouve k , et qu'ils sont disjoints, alors on est certain que toute grille partielle conduisant à G comportera au moins une donnée dans chaque ensemble inévitable et donc en comportera au moins k . Lorsque certains ensembles inévitables ont une intersection commune, la situation, un peu plus difficile, impose tout de même un nombre minimum de données à toute grille partielle conduisant à G.

La lutte finale

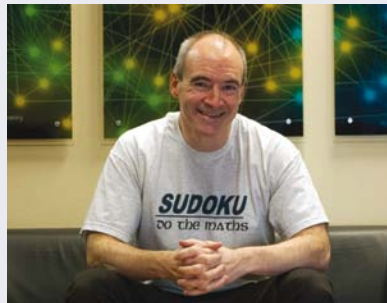
Le travail pour traiter une grille complète G et montrer qu'il faut toujours au moins 17 données se ramène alors à la recherche et à l'exploitation d'ensembles inévitables aussi nombreux que possible. Après divers essais, les chercheurs ont adopté la stratégie suivante, qui a été appliquée à chacune des 5,5 milliards de grilles complètes du catalogue de G. Fowler.

(a) Déterminer (en mettant au point les algorithmes nécessaires) tous les ensembles inévitables de moins de 13 cases. Trouver tous les ensembles inévitables sans limitation de taille, ce qui *a priori* semble meilleur, conduit à plus de calculs au total.

(b) En déduire (toujours en mettant soigneusement au point les algorithmes nécessaires) toutes les sous-grilles de 16 données convenables pour cette famille d'ensembles inévitables, c'est-à-dire ayant au moins un élément dans chaque ensemble inévitable. Cette partie retient considérablement moins de grilles partielles à 16 éléments que les $3,4 \times 10^{16}$ grilles partielles *a priori* envisageables. En moyenne, une grille complète de sudoku a 360 ensembles inévitables de moins de 13 cases.

(c) Pour les grilles partielles retenues, utiliser un programme de test qui indique si oui ou non il y a une solution unique.

6	3	9	2	4	1	7	8	5
2	8	4	7	6	5	1	9	3
5	1	7	9	8	3	6	2	4
1	2	3	8	5	7	9	4	6
7	9	6	4	3	2	8	5	1
4	5	8	6	1	9	2	3	7
3	4	2	1	7	8	5	6	9
8	6	1	5	9	4	3	7	2
9	7	5	3	2	6	4	1	8



LE MATHÉMATICIEN GARY MCGUIRE (ci-dessus) et un sudoku ayant 29 sous-grilles correctes à 17 données (ci-contre).