

Le projet de généalogie mathématique a 21 ans

Colm Mulcahy

Répondre à un besoin (et avoir besoin d'une adresse)

Le Mathematics Genealogy Project (MGP) fournit en ligne des informations sur les directeurs de thèse et les descendants mathématiques de plus de 200 000 mathématicien(ne)s, une bonne partie des 825 000 auteurs sur MathSciNet et bien plus que les 2 700 biographies de l'histoire des mathématiques de MacTutor. Il n'en a pas toujours été ainsi.

En juin 1984, le département de mathématiques de l'Université de Chicago organisa une conférence marquant le départ à la retraite officiel d'Irving Kaplansky (1917-2006), l'algébriste canadien qui avait été parmi les premiers à réussir brillamment l'examen Putnam en 1938. Kap, comme on le surnommait, était loin d'être prêt à prendre sa retraite : il s'apprêtait à prendre la direction du MSRI en tant que deuxième directeur.

Lors de la réunion de Chicago, on a beaucoup insisté sur le fait que l'invité d'honneur avait supervisé 55 doctorants, dont Arlen Brown, Hyman Bass, Donald Ornstein, Joe Rotman, Judy Sally et mon propre directeur de thèse, Alex Rosenberg.

Si je me souviens bien, une liste de documents a circulé, recensant ces personnes, ainsi que leurs doctorants, etc., afin de démontrer la fécondité intellectuelle de Kaplansky. Ironie du sort, le propre conseiller de Kap, Saunders Mac Lane, était également présent à cette réunion. Il a simplement ri lorsque je l'ai interpellé dans un couloir et lui ai fait remarquer qu'une liste de tous ses descendants serait encore plus impressionnante.

Aujourd'hui, contrairement à il y a trente ans, il est facile de consulter ses ancêtres et descendants sur MGP en ligne¹, comme l'illustre partiellement la figure 1. Colm Mulcahy, professeur de mathématiques au Spelman College d'Atlanta, y montre que Kaplansky fut le premier des 42 doctorants supervisés par Mac Lane. Il fut également le plus prolifique d'entre eux, avec une descendance qui compte aujourd'hui plus de 850 personnes. Parmi les doctorants de Mac Lane figurent le logicien Anil Nerode, le théoricien des groupes John Thompson, le théoricien des ensembles Robert Solovay et l'algébriste David Eisenbud, chacun ayant plus de 100 descendants. Les notices de Kaplansky et de Mac Lane renvoient à leurs fiches MacTutor et MathSciNet, une caractéristique commune à de nombreuses notices importantes.

Comme pour tout arbre généalogique, il existe des complications. Par exemple, Eisenbud avait deux directeurs de thèse. Mac Lane lui-même aussi : Hermann Weyl et Paul Bernays. La lignée Weyl remonte, par Hilbert et Lindemann, jusqu'à Klein, qui avait également deux directeurs de thèse.

Communiqué par Steven J. Miller

Transcription en L^AT_EX et traduction via outils google : Denise Vella-Chemla, novembre 2025.

On se reportera aux figures de l'article à l'url :

<https://www.ams.org/publications/journals/notices/201705/rnoti-p466.pdf>.

Note de la traductrice : il est vraisemblablement sous-entendu dans le titre que le projet a en quelque sorte atteint la majorité.

En suivant ces deux lignées, on rencontre d'autres figures emblématiques des mathématiques : Lipschitz, Dirichlet, Poisson et Fourier, Lagrange et Laplace, Euler, d'Alembert, Plücker et Gauss. Bien sûr, Gauss a beaucoup à se faire pardonner : le MGP estime actuellement le nombre de ses descendants à environ 78 000.

Origines de l'espèce

Le Mathematics Genealogy Project a été lancé il y a vingt et un ans, en 1996, par Harry Coonce, alors mathématicien à l'université d'État de Mankato, dans le Minnesota. Il était motivé en partie par la difficulté qu'il avait rencontrée en essayant de déterminer l'identité de son propre "grand-père mathématique", c'est-à-dire le directeur de thèse de son directeur de thèse. Il s'agissait d'Einar Hille, mais à l'époque, il n'existait aucun moyen de le vérifier, ni en bibliothèque ni sur internet. (Bruce Reznick se souvient que Constance Reid, dans son ouvrage *The Search for E. T. Bell*, indiquait ne pas avoir trouvé le nom de son directeur de thèse). Ce printemps-là, Harry écrivit à des centaines de départements et, au moment du lancement du site web original du MGP en septembre, il avait recensé 3 500 docteurs.

La figure 1 montre Les ancêtres les plus récents d'Irving Kaplansky issus du Mathematics Genealogy Project ; l'arbre généalogique complet – dont seule une partie est présentée ici – compte désormais 850 descendants et remonte à plus de sept siècles.

En un an, le projet avait atteint 10 000 entrées. Lorsque l'article "A Labor of Love : The Mathematics Genealogy Project" d'Allyn Jackson a paru dans les Notices en 2007, le MGP comptait plus de 100 000 personnes. À cette époque, Coonce avait pris sa retraite, s'était installé à l'Université d'état du Dakota Nord et avait accueilli un assistant étudiant qui a finalement repris le projet : Mitch Keller, aujourd'hui mathématicien à l'Université Washington et Lee de Lexington, en Virginie.

De ses débuts difficiles, avec peu ou pas de financement, le projet est devenu un pilier de la communauté mathématique. Ce programme est désormais financé par le Clay Institute et l'AMS, et durant l'été 2016, le nombre de thèses soumises au MGP a franchi une nouvelle étape importante, dépassant les 200 000. La croissance a été globalement linéaire, comme le montre la figure 2.

Bien que MGP collecte des informations classées sous différentes appellations telles que Ph.D., D.Phil. ou D.Sc., aucune distinction de ce type n'est faite dans cet article. Le premier doctorat américain en mathématiques (obtenu à Yale en 1861) a été décerné à Arthur Wright ; ses travaux portaient sur la mécanique des satellites. Comme plus de 75 % des personnes répertoriées dans MGP, moi y compris, Arthur n'a eu aucune descendance. La nature du travail doctoral et le système de direction de thèse sont abordés. La nature du travail doctoral et le système d'encadrement - sans mentionner le fait qu'au jour d'aujourd'hui, l'obtention d'une thèse est un prérequis pour pouvoir avoir une existence académique - est sans doute issu d'anciens modèles français et allemands, et tout cela a mis du temps à s'implanter.

Un principe inclusif

Le principe de la base de données MGP repose sur une philosophie très inclusive, accueillant les

candidatures en didactique des mathématiques, en statistique, en informatique ou en recherche opérationnelle. Les domaines de thèse sont sélectionnés à l’aide d’une liste déroulante de la Classification mathématique par sujets de 2010. Ainsi, en théorie, toutes les thèses en physique théorique, géophysique, biologie, astronomie, statistique, informatique et/ou recherche opérationnelle pourraient être retenues. On compte déjà environ 2 000 thèses classées dans les catégories “Théorie des systèmes, contrôle” et “Biologie et autres sciences naturelles” ; environ 3 400 dans la catégorie “Recherche opérationnelle, programmation mathématique” ; et plus de 18 000 dans la catégorie “Informatique”. La page “Extrema” du site répertorie les 75 “meilleurs” directeurs de thèse (par ordre décroissant), qui surpassent tous Kaplansky, même si plusieurs d’entre eux soulèvent des interrogations. En tête de liste figure l’ingénieur électricien chinois C.-C. Jay Kuo, dont le doctorat du MIT est classé dans la catégorie “Information et communication, circuits”. Au cours des vingt-cinq dernières années à l’Université de Californie du Sud, il a supervisé 137 thèses de doctorat. En deuxième position, on trouve l’analyste numérique français Roger Temam avec 119 étudiant(e)s depuis 1972, et en troisième position, l’économiste et informaticien Andrew Whinston (dont le nombre d’Erdős est 2), qui a encadré 104 étudiant(e)s depuis 1970. Kolmogorov, en treizième position avec 82 doctorants, est sans doute le nom le plus connu de la plupart des mathématicien(ne)s. Hilbert (David) a eu 79 étudiant(e)s, Blackwell (David Harold) 65 et Stanley (Richard) 60. Avec 55, Kaplansky (Irving) se classe 87e.

La figure 2 illustre la croissance du nombre de fiches de la base de données MGP.

Les enfants représentés et leur valeur relative

Les particuliers, comme les départements, intéressés par leur patrimoine, apprécient de voir les moments marquants mis en valeur dans des arbres généalogiques. MGP propose des affiches de généalogie mathématique personnalisées (payantes). Mitch remarque : “Les affiches individuelles sont très appréciées pour les anniversaires, les départs à la retraite et comme cadeaux de fin d’études.”. Stephan Garcia, du Pomona College, confirme : “Nous avons acheté une grande affiche MGP pour tout le département il y a quelques semaines.” (Voir figure 3). De même, Michael Wolf de Rice déclare : “Lorsque j’étais directeur de département, j’ai acheté une version du MGP afin de présenter visuellement aux visiteurs le patrimoine intellectuel du département.”. Mais qui utilise le MGP, outre la simple consultation ou le plaisir de se vanter de ses ancêtres ? Certainement les contributeurs de Wikipédia : nombre de pages Wikipédia consacrées aux mathématicien(ne)s de l’ère moderne contiennent un lien externe vers l’entrée MGP correspondante, et certaines informations proviennent de ce lien. Il y a aussi les rédacteurs de revues scientifiques. Norman Richert, rédacteur en chef de Mathematical Reviews, siège au comité consultatif du MGP. Il partage l’avis des autres rédacteurs quant à l’utilité du MGP pour “dresser un portrait des personnes et de leur place dans le paysage mathématique”, ainsi que pour éviter d’éventuels conflits d’intérêts. Il soulève un autre point : “Lorsque je pense aux auteurs et que je les invite à devenir relecteurs, je tiens à savoir s’ils sont titulaires d’un diplôme (de nombreux doctorants publient aujourd’hui davantage d’articles pendant leurs études qu’auparavant)”. Les liens MGP sur les pages de profil des auteurs dans MathSciNet répondent à cette question avec une forte probabilité (mais pas à 100 %).

Janet Beery (Université de Redlands) et Carol Mead (archiviste des Archives des mathématiques

américaines à l'Université du Texas à Austin) ont utilisé MGP pour rassembler des informations pour leur article de Convergence intitulé "Qui est ce(tte) mathématicien(ne) ? Images de la collection photographique Paul R. Halmos" (plus de soixante pages web). Carol explique : "C'est un outil de référence formidable. Je l'utilise pour identifier ou vérifier l'identité d'une personne. Et, n'étant pas mathématicienne, il me permet d'en apprendre davantage sur des mathématicien(ne)s que je connais moins.". Lev Borisov, de l'Université Rutgers, ajoute : "Je pense que ce sera un outil précieux pour les historiens des mathématiques. J'imagine également qu'il sera utile aux spécialistes des sciences sociales et aux agences gouvernementales."

Naissance, filiation et décès

Les trois piliers de la généalogie traditionnelle sont la documentation des naissances, des filiations et des décès, et les versions académiques des deux premières se retrouvent dans le MGP. Cependant, bien que les détails de la vie et du décès n'y soient pas enregistrés, le site s'avère utile pour les rédacteurs de nécrologies, tout comme les arbres généalogiques classiques. "Lorsqu'un(e) mathématicien(ne) décède, nous aimons indiquer où il a obtenu son doctorat, la date d'obtention et parfois le nom de son directeur de thèse", explique Mike Breen, responsable de la communication de l'AMS, qui a rédigé de nombreux articles "À la mémoire de..." sur le site web de l'AMS. "Nous aimons mentionner le directeur de thèse lorsqu'il s'agit d'une personnalité connue. Je pense que cela permet au lecteur de mieux comprendre le parcours de la personne, tout comme le fait le MGP lui-même. Il peut être intéressant de mentionner par exemple que telle personne est apparentée à Kolmogorov par deux générations de mathématicien(nes)".

Norman Richert déclare à nouveau : "La valeur de MGP pour la communauté est immense. La recherche en mathématiques a toujours impliqué un réseau complexe de personnes. Le graphe de collaboration de MathSciNet en donne un aperçu, tout comme le graphe directeur de thèse/thésard(e) de MGP. Les mathématicien(ne)s s'intéressent depuis longtemps à ces interactions, peut-être plus que les scientifiques de toute autre science. À ses débuts, certains semblaient penser qu'il s'agissait d'un passe-temps futile. La plupart des gens, je crois, en ont aujourd'hui compris l'importance ; l'AMS, en tout cas, l'a bien compris."

Limites

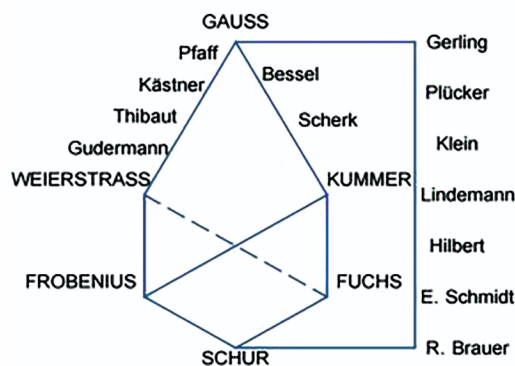
Bien entendu, l'utilité de MGP est limitée. Elle ne permet de remonter que jusqu'à une certaine période et ne prend en compte qu'un certain type de formation mathématique. À l'instar de la généalogie traditionnelle, MGP tend également à ignorer les descendances non déclarées. Nombreux sont ceux qui considèrent que leur véritable mentor mathématique est quelqu'un d'autre que leur supérieur hiérarchique officiel. Comme dans toute famille, "parents" et "enfants" ne s'accordent pas toujours sur la nature exacte de leur lien de parenté. Plusieurs raisons, d'ordre institutionnel ou personnel, peuvent expliquer ce décalage entre les informations officielles et la réalité. Dans la plupart des cas, ces anomalies restent non documentées en ligne. Par ailleurs, les jeunes s'intéressent généralement moins à la généalogie en général que leurs aînés et ne sont pas toujours réceptifs aux incitations à inscrire leurs données dans la base MGP. Comme pour toutes les archives des sciences, technologies, ingénierie et mathématiques (STEM), il est important que les groupes sous-représentés en mathématiques soient aussi présents que possible dans le MGP, même si leur statut minoritaire

n'est pas visible au premier abord : aucune information d'âge, de sexe, de nationalité, de photo ou autre que le nom n'est indiquée pour les fiches. (Le drapeau du pays affiché pour chaque diplôme correspond à la désignation actuelle du pays de l'établissement qui a délivré le diplôme). Pour ces seules raisons, il est particulièrement important que les directeurs de thèse et les départements soumettent, corrigent et mettent à jour régulièrement les informations dans la base de données MGP. Enfin, se pose la question de l'exactitude : n'importe qui peut soumettre des données dans la base MGP, et dans la plupart des cas, elles sont acceptées sans vérification préalable. Il existe certainement des cas où l'on suppose que X a été étudiant(e) de Y "avec une forte probabilité", même si cela ne peut être prouvé ; parfois, ni les thèses en question ni les documents officiels correspondants ne subsistent.

Structure et support

Les personnes curieuses de connaître la structure du graphe formé par MGP peuvent trouver des informations sur le site web. Vers 2008, Bud Brown de Virginia Tech a remarqué que le graphe MGP n'était pas planaire et a dessiné le sous-graphe de la figure 4 illustrant ce fait.

En juillet 2016, environ 90 % des entrées de MGP formaient une seule composante connexe, dont environ 75 % étaient occupés par quelque 7 300 sommets racines (personnes sans directeur de thèse répertorié) et quelque 137 000 feuilles (personnes sans étudiant(e) répertorié(e)). On trouve également des sommets isolés, comme celui du mathématicien irlandais William Rowan Hamilton. Il semble avoir été inclus par simple courtoisie : il n'avait pas de doctorat (ni de descendants connus). De manière problématique, son entrée, dépourvue de mention, comporte les mentions "Directeur de thèse : Inconnu" et "Aucun(e) étudiant(e) connu(e)", la première suggérant qu'il suffirait peut-être d'approfondir les recherches. La désignation "Aucun" serait plus utile dans les deux cas, et en général, il serait souhaitable d'avoir un moyen de distinguer entre les listes (éventuellement vides) de descendants académiques connus pour être complets et celles pour lesquelles des données supplémentaires pourraient éventuellement être ajoutées ultérieurement.



La figure 4 est une preuve sans mot du fait que le graphe représentant les liens entre fiches de la base de données MGP est non planaire, puisqu'il contient un sous-graphe homéomorphe à $K_{3,3}$ reliant Gauss, Frobenius, et Fuchs à Weierstrass, Kummer, et Schur.

Certains départements de mathématiques offrent un soutien structuré (implicite) au MGP en affichant fièrement en ligne des listes de tou(te)s leurs étudiant(e)s diplômé(e)s ayant réussi, remontant aux premiers temps, deux exemples notables étant ceux de l'Université de Colombie-Britannique et de l'Université College Cork (UCC) en Irlande. Bien sûr, cela laisse à d'autres le travail non trivial

de soumettre les informations pertinentes (manquantes) sur les thèses au MGP, et dans le cas de l'UCC, votre serviteur s'y est essayé.

La plupart des sites web universitaires proposent des moteurs de recherche permettant de suivre l'existence (et parfois aussi le texte) des thèses, bien que cette fonctionnalité soit souvent limitée aux diplômes les plus récents. Malheureusement, beaucoup ne permettent pas de rechercher par sujet ou par directeur de thèse ; certaines institutions semblent même ne conserver aucune trace de ces informations. Le site de Caltech¹ est l'un des meilleurs, puisqu'il permet même de rechercher par directeur de thèse ou par membre du comité. Il pourrait être intéressant pour les étudiant(e)s de ces institutions de veiller à ce que toutes les thèses pertinentes soient ajoutées au MGP. Si le MGP dépend entièrement des contributions volontaires pour s'enrichir et gagner en précision, il a également besoin d'un soutien financier. Le directeur général, Mitch Keller, souligne : "Tous les membres du MGP sont reconnaissants du soutien apporté par la communauté mathématique au cours des vingt dernières années, et nous espérons continuer à être une ressource précieuse pour de nombreuses années encore.". Le MGP utilise depuis longtemps les dons pour employer des étudiant(e)s chargés du traitement des données. Depuis son installation à la NDSU en 2002, il s'agit principalement d'assistants de recherche qui effectuent ce travail en plus de leurs cours. Le budget annuel, depuis 2008-2009, oscille entre 6 000 et 9 000 dollars, dont environ la moitié est consacrée aux salaires des assistants de recherche. Le reste sert généralement aux frais de déplacement et à l'impression d'affiches. Nous nous efforçons d'accroître le montant des dons reçus chaque année afin d'améliorer le projet. Actuellement, notre délai d'attente est d'un mois en raison d'un manque de personnel, principalement dû à des problèmes de financement, et nous devons déjà anticiper les coûts de notre participation à l'ICM en 2018. Les dons destinés à soutenir le travail du MGP sont les bienvenus, qu'il s'agisse de dons ponctuels (y compris un legs) ou de dons réguliers. Vous pouvez contribuer en cliquant sur le bouton "Faire un don" dans le menu latéral de la page d'accueil.

Colm Mulcahy du Lycée Spelman d'Atlanta, est l'auteur du livre *Mathematical Card Magic : Fifty-Two New Effects* (A. K. Peters/CRC Press, 2013), vice-président de *Gathering 4 Gardner*, et le créateur et conservateur des *Annals of Irish Mathematics and Mathematicians* en ligne.

1. Le lien "thesis.library.caltech.edu" ne semble pas fonctionner en 2025.