01

MATHÉMATIQUES ET OUTILS DE MODÉLISATION

Jean-Luc SAUVAGEOT
Président

Jean-Paul Allouche
Michèle Audin
Gérard Besson
Michel Brion
Jérome Buzzi
Claire Chainais
Fabienne Comte
Pierre-Olivier Flavigny
Maurice Galeski
Patrick Gerard
Gilles Godefroy
Florence Hubert
Michel Langlais
Christine Noot-Huyghe
Michel Pierre
Geoffrey Powell
Pierre Torasso
Annie Touchant
Frank Wagner
Wendelin Werner

L'évènement qui est incontestablement le plus important pour la communauté mathématique française est l'attribution de la Médaille Fields (l'équivalent du prix Nobel) à Laurent Lafforgue lors du congrès internationnal de Pékin (Août 2002). Laurent Lafforgue est professeur permanent à l'Institut des Hautes Études Scientifiques (IHES) et a fait une grande partie de sa courte carrière au CNRS. Deux autres évènements, plus dramatiques, ont marqué cet été exceptionnel : il s'agit de la disparition de Laurent Schwartz, en juillet 2002, fondateur de la célèbre théorie des distributions (médaille Fields 1950) et de celle de René Thom, en septembre 2002, qui donna à la topologie différentielle ses lettres de noblesse (médaille Fields 1958). Comme une passation de témoin, la conjonction de ces évènements atteste de la continuité (il serait incongru de parler de vitalité) dans l'excellence des mathématiques françaises. Par les quelques exemples qui suivent nous espérons mettre en évidence cette qualité et le rôle essentiel joué par le CNRS.

INTRODUCTION

Le premier partenaire du CNRS, en mathématiques, est l'Université. La quasi-totalité de la communauté universitaire, près de 3 000 enseignants-chercheurs, est structurée dans les unités associées au CNRS, autour d'une section unique de 350 chercheurs. Au prix d'un investissement en personnels, chercheurs et ITA, qui demeure relativement modeste, le CNRS se voit assigner un rôle moteur dans cette communauté très active sur le plan de la recherche. Pourtant, ce rapport se focalisera essentiellement sur l'ouverture des mathématiques, à la fois vers les autres disciplines et vers les applications industrielles.

La Mathématique occupe une place singulière dans l'ensemble de la culture scientifique. Par son ancienneté peut-être, mais surtout par sa présence dans toutes les Sciences exactes (et même plus) dont elle est d'une certaine manière le langage universel. Le début du siècle précédent a vu apparaître en France un phénomène lourd de conséquences pour notre communauté : il s'agit de la séparation entre des mathématiques qualifiées de pures et celles dites appliquées, rupture qui commence à peine à être étudiée, voire comprise, par les historiens. Le retour à une perception plus unitaire de la culture mathématique, c'est-à-dire à un continuum entre les aspects les plus fondamentaux, l'ouverture vers les autres disciplines, et les applications industrielles, est sensiblement amorcé aujourd'hui. Il accompagne peut être l'apparition d'une conception moins romantique de notre discipline, ou d'une nouvelle génération de mathématiciens, mais il coïncide surtout avec la prise en compte d'un phénomène passionnant, et spectaculaire : de plus en plus de mathématiques très théoriques trouvent des applications inattendues, et ceci rapidement. Le passage de concepts très abstraits à des conséquences industrielles et commerciales ne requiert plus nécessairement le filtre des mathématiques appliquées ; les exemples les plus frappants étant donnés par la cryptographie et la mathématique du génome. Parallèlement, les mathématiques appliquées ont acquis une efficacité remarquable. Tout ceci contribue à donner à notre discipline une fraîcheur et une souplesse sans doute plus perceptibles que par le passé.

Le texte qui suit n'a pour seul but que de donner des clés pour comprendre ce que les Mathématiques représentent dans notre pays, surtout en termes de transferts de la théorie vers les applications : ce n'est certainement pas leur seul aspect, mais c'est celui le plus souvent sous-estimé par les décideurs. Nous devons insister sur le rôle du CNRS : la France est le deuxième pays au monde pour cette discipline (peut-être le premier si on se rapporte au nombre d'habitants) et le CNRS est une des raisons de ce succès : non que l'essentiel des avancées scientifiques soit produit par les seuls chercheurs de cet organisme, mais par l'importance de son activité d'organisation. La section 01, comme nous l'avons déjà dit, est la seule à pouvoir se prévaloir de fédérer l'ensemble de la communauté française dont elle a la charge.

1 – MATHÉMATIQUES FONDAMENTALES

La vitalité des mathématiques fondamentales françaises est un phénomène ancien ; tous les aspects de cette discipline sont concernés : analyse, géométrie et topologie, algèbre, comme en témoigne les trois personnalités citées en prologue de ce texte. Pour rester proche de l'actualité (la Médaille Fields de Laurent Lafforgue) nous nous contentons d'esquisser quelques épisodes de la passionnante histoire du programme de Langlands (mathématicien canadien en poste à l'université de Yale).

Si l'origine de l'aventure peut être située dans la Grèce antique, avec les tentatives de résolutions des équations dites Diophantiennes (équations à coefficients entiers dont on cherche des solutions sous formes de nombres entiers), c'est vraiment le résultat de C. F. Gauss, appelé réciprocité quadratique, qui en est le point de départ à l'époque moderne (XIX^{e} siècle). Il s'agit d'une condition portant sur deux nombres premiers p et q pour que p soit un carré modulo q, c'est-à-dire pour qu'il existe un entier x tel que $x2 - p$ soit divisible par q. L'importance et la profondeur de ce résultat ont poussé les chercheurs à tenter d'établir des lois de réciprocité pour des degrés plus généraux (c'est-à-dire pour des équations plus générales que $x2 - p = 0$ [modulo *p*]) et aussi pour des ensembles de nombres plus généraux que les nombres entiers ou rationnels.

Ces ensembles qui possèdent les quatre opérations, comme c'est le cas pour les nombres réels, sont appelés « corps ». La théorie de Galois, du nom du mathématicien français Evariste Galois, traite justement de la notion d'extension de corps. Plus précisément, la théorie de galois associe à tout corps de nombres E (extension finie du corps des nombres rationnels) un groupe GE, de sorte que si le corps F contient le corps E, alors GF est inclus dans GE ; le groupe de Galois de l'extension est alors GE/GF. E. Galois utilisa cette splendide théorie pour déterminer les équations résolubles par radicaux. La théorie du corps de classes qui occupa les mathématiciens de la fin du XIX^{e} siècle jusqu'à la moitié du XX^{e} siècle a permis de classifier les extensions abéliennes (commutatives) ; dans cette perspective, Emil Artin établit en 1927 une loi de réciprocité générale et abstraite qui englobait toutes celles connues précédemment (dont celle de Gauss). C'est dans l'espoir de généraliser cette théorie que Robert Langlands mis au point son programme qui est un tissu de conjectures établissant un lien entre les représentations de dimensions finies du groupe de Galois attaché à l'extension d'un corps K (corps de nombres ou de fonctions) et certaines représentations, dites automorphes, associées au groupe linéaire GL (n, K). La théorie des représentations automorphes peut être vue comme une extension de la théorie de Fourier, elle-même élaborée pour étudier les phénomènes vibratoires (propagation de la chaleur et des ondes). La difficulté des conjectures de Langlands se mesure par l'entier n, le cas $n = 1$ correspondant à la théorie du corps de classes. De nombreux sous-cas peuvent être envisagés suivant les valeurs de n et le type de corps considéré. En 1994, Andrew Wiles établit un cas particulier pour les corps de nombres, pour $n = 2$, et en déduit le théorème de Fermat. Au cours des années 70 Vladimir Drinfeld s'attaqua au cas des corps de fonctions et établit la correspondance de Langlands pour $n = 2$ pour ces corps. Pour ce faire il introduit de nouveaux objets géométriques abstraits appelés les chtoucas de Drinfeld (signalons également la notion de module de Drinfeld, généralisation des courbes elliptiques et qui joue déjà un rôle en cryptologie). Il obtient la médaille Fields en 1990.

Laurent Lafforgue, s'appuyant sur les constructions de V. Drinfeld, démontre la correspondance de Langlands pour les corps de fonctions et tous les rangs $n \geq 1$. La preuve est d'une abstraction et d'une virtuosité technique époustouflante. Elle utilise aussi la vision géométrique et la généralisation partielle de la théorie de Galois d'Alexander Grothendieck (mathématicien français, médaille Fields 1966 ; il fut l'élève de Laurent Schwartz), qui engagea la refonte de la géométrie algébrique après des travaux pionners de Jean-Pierre Serre (médaille Fields 1954).

Cette aventure, qui est loin d'être terminée, est significative du fonctionnement des mathématiques, et de la qualité et de l'influence de l'école française.

2 – ANALYSE NUMÉRIQUE

La discipline que l'on appelle traditionnellement l'analyse numérique combine en réalité de multiples facettes. Elle consiste à étudier les représentations mathématiques d'une réalité extérieure (généralement physique ou mécanique, mais de plus en plus biologique, économique ou relevant des sciences humaines, comme le comportement psychologique par exemple) en la modélisant à l'aide d'équations différentielles ou d'équations aux dérivées partielles. L'étude mathématique des équations obtenues (indispensable pour décider de la qualité du modèle obtenu), leur discrétisation (c'est-à-dire leur traduction en instructions assimilables par un ordinateur) et la réalisation des programmes informatiques correspondants constitue l'éventail des activités qui lui sont dévolues. De fait, son activité est à l'interface entre la discipline dont relève l'objet étudié, les mathématiques traditionnelles et l'informatique.

L'analyse numérique interagit potentiellement avec toutes les autres sciences.

Les activités dans les domaines aéronautique, spatial, nucléaire et militaire ont constitué le terreau de son développement en France au cours des décennies 1960-1980. Depuis, un foisonnement de développements dans de nouveaux secteurs s'est fait jour :

– électromagnétisme, télécommunications (électromagnétisme, optique) ;

– technologies spatiales (combustion, plasmas) ;

– chimie (agrégats, pharmaco-synthèse) ;

– biologie (biomécaniques, épidémiologie, immunologie, neurologie, etc.) ;

– électronique ;

– informatique (algorithmique, réalité virtuelle) ;

– nouvelles technologies de la communication (fibres optiques, communication sans fil, compatibilité électromagnétique, électronique, etc.) ;

– génie rural (gestion de l'eau, des déchets, retraitement), prédiction des risques naturels, etc.

Transversalement à ces applications, on assiste actuellement à un développement intensifs des travaux sur les couplages de modèles (couplage fluide-structure, couplage océan-atmosphère, couplages d'échelles microscopique-mésoscopique, etc.), développement rendu nécessaire par l'exigence d'applications comme celles à la météorologie ou aux technologies spatiales par exemple. Ces nouvelles approches bouleversent les fondements traditionnels de l'analyse numérique que sont les méthodes d'approximation par éléments finis ou différences finies, et favorise la forte montée en puissance des méthodes sans grille (méthodes particulaires), sans structures de connectivité rigides (volumes finis, méthodes boîtes) ou multiéchelles (ondelettes), ainsi qu'à un renouveau des méthodes traditionnelles (éléments finis joints).

L'exigence de qualité des industriels a engendré le développement de méthodes de contrôle des approximations numériques (estimations a posteriori) et la mise au point de critères de raffinement de maillages, essentiels pour capturer les échelles multiples des phénomènes physiques. L'approche des phénomènes multiéchelles a également dynamisé les études asymptotiques, notamment dans le domaine de l'homogénéisation, technique particulièrement prisée en science des matériaux, ou pour la modélisation des gaz, plasmas et semiconducteurs.

Enfin, on ne se contente plus de chercher à prévoir un phénomène. On cherche à concevoir un outil répondant à des critères précis de performances et de coûts. Dans cette optique, le contrôle et l'optimisation (notamment l'optimisation de forme) ont vu se développer ces dernières années des outils puissants (l'optimisation topologique, la différenciation automatique par programme) qui révolutionnent les approches traditionnelles. Celles-ci sont relayées par les progrès considérables effectués dans les techniques de résolution de grands systèmes linéaires (méthodes de type Krylov) qui, associées à la parallélisation massive, mettent à la disposition de l'industriel des capacités de calcul inégalées. Il s'agit là d'une interface évidente avec les STIC puisque, par tradition, ces disciplines sont à l'interface des mathématiques, de l'informatique et de l'automatique.

3 – INTERACTION DES MATHÉMATIQUES DISCRÈTES, DE LA THÉORIE DES NOMBRES ET DE LA LOGIQUE AVEC L'INFORMATIQUE ET LA PHYSIQUE : QUELQUES EXEMPLES

Les mathématiques discrètes, la théorie des nombres et la logique entretiennent de nombreux liens avec l'informatique (théorique en particulier). Citons d'abord les suites engendrées par automates finis dont les propriétés arithmétiques et dynamiques sont étudiées par les théoriciens des nombres et les ergodiciens (depuis les travaux de Keane en 1968, de Cobham de 1967 à 1972, de Christol en 1979 et de Christol, Kamae, Mendès France et Rauzy en 1980) : ces suites permettent d'éclairer d'un jour nouveau des questions de théorie des langages (langages réguliers, automates finis à fonction de sortie, combinatoire des mots) ou des questions d'engendrement de suites « quasi-aléatoires ». Citons d'autre part un exemple concernant la logique : la théorie de la démonstration peut être identifiée à l'étude des langages de programmation grâce à la correspondance de Curry-Howard (formulée dans les années 50) entre la logique intuitionniste et le lambda-calcul typé. Ces deux directions de recherche sont très actives et de nombreux articles contemporains leur sont consacrés (pour doner un seul élément d'appréciation, une nouvelle rubrique dans la classification de l'AMS a été créée récemment : 11B85 « automata sequences »). En physique, les suites automatiques et la combinatoire des mots apparaissent d'abord dans l'étude de modèles résolubles comme le modèle d'Ising. Puis la découverte des quasicristaux (Shechtman, Blech, Gratias et Cahn en 1984) et leurs liens avec les pavages de Penrose (1972) conduisent les mathématiciens, les informaticiens et les physiciens statisticiens à étudier dans ce cadre de nombreuses familles de suites « complexes à engendrement simple » (suites sturmiennes, suites automatiques etc.).

Des interactions ou partenariats existent tant au CNRS qu'à l'université : depuis l'époque du PRC math.-info. jusqu'aux contacts et coopérations formels ou informels avec des collègues de laboratoires d'informatique, par exemple de l'ancien LITP – Laboratoire d'informatique théorique et de programmation, à Jussieu – devenu le LIAFA Laboratoire d'informatique Algorithmique, Fondements et Applications. On peut aussi noter que des chercheurs en 01 sont affectés dans des laboratoires en section 07 – ou l'inverse. Enfin il existe des séminaires communs ou à audience mixte (par exemple à l'INRIA, ou entre le LRI et l'école Polytechnique) et il y a de nombreux contacts avec des collègues physiciens théoriciens (physique statistique) du CEA.

La valorisation des travaux dans ce domaine va essentiellement dans trois directions :

– la théorie de la démonstration et les langages de programmation ainsi que les techniques de vérifications de programmes ;

– suites engendrées par automates finis et génération aléatoire ;

– applications indirectes des suites automatiques dans les quasi-cristaux.

4 – STATISTIQUES, TÉLÉCOMMUNICATIONS ET TRAITEMENT DU SIGNAL

Les mathématiques et la physique théorique permettent de mesurer l'information de manière quantitative à travers la notion d'entropie. Cette quantité fournit, en mécanique statistique, une mesure du désordre d'un système. Dans la théorie de l'information elle détermine combien d'éléments binaires sont nécessaires pour transmettre, après compression, un signal modélisé par un processus aléatoire.

En bref, il s'agit de construire et d'étudier des méthodes de codage qui permettent soit une compression optimale de sources de plus en plus variées, soit la transmission la plus fiable possible sur un canal imparfait, qui introduit des erreurs. Les applications à la théorie de la prédiction sont très intéressantes, car la compression d'une source peut être interprétée comme une façon de prédire le mieux possible son comportement futur à partir de son comportement passé.

C'est dans le domaine de la prédiction que la théorie de l'information rencontre la statistique, autre élément important des STIC. L'objet en est de deviner le plus exactement possible quelle loi de probabilité a généré un ensemble de données observées qui présente des fluctuations aléatoires, soit par nature, soit parce que l'expérimentateur qui a généré ces données a introduit lui même cet aléa pour se protéger de certains biais (sondages, plans d'expérience). La collecte de données de plus en plus nombreuses et de plus en plus variées a provoqué une explosion de la statistique. Les statisticiens se posent la question de l'analyse des données à partir de modèles imparfaits ou imparfaitement connus. Les techniques de sélection des modèles utilisent des résultats théoriques (en probabilités par exemple) très récents, et sophistiqués, qui permettent de contrôler le risque.

De nouveaux problèmes sont apparus dans les communautés statistiques et informatiques. L'étude des modèles, des algorithmes et la compréhension de leur comportement sont un défi d'actualité, tant au niveau théorique qu'au niveau des applications à la fouille des données, la fouille des textes. Des applications, par exemple à la bio-informatique, introduisent la problématique de l'apprentissage de séquences. On assiste à un fort développement des activités de recherche sur ces thèmes dans le monde. Un modèle pour être efficace doit couvrir un grand nombre de cas à l'aide d'un petit nombre de paramètres.

Dans le cas de modèles fonctionnels, cette propriété est étudiée par la théorie de l'approximation, qui entretient des liens étroits avec l'analyse harmonique.

Un exemple très célèbre, qui a son origine en France, est la théorie des ondelettes et ses liens avec l'approximation dans les espaces de Besov, développée à l'origine par Y. Meyer. Des problèmes passionnants restent ouverts dans le domaine de la représentation des signaux multi-dimensionnels. Dans le domaine des applications, la compression du signal vidéo reste un défi.

Les télécommunications posent aujourd'hui aux mathématiciens des problèmes très ardus et d'une importance pratique capitale ; citons, à titre d'exemples, les problèmes de dimensionnement et de reconfiguration, la compréhension théorique des turbo-codes, la gestion des réseaux. Il en est de même du traitement du signal avec les problèmes de codage d'images fixes et animées et la reconnaissance de la parole.

Bien que la France ait joué un rôle de précurseur dans certains des aspects développés ci-dessus, elle accumule un retard important. Le CNRS doit s'engager de manière ferme et volontariste dans ce qui est un des principaux défis technologiques de l'avenir. Les retombées économiques sont considérables.

Trois pôles sont impliqués : la physique, les mathématiques et l'ensemble des STIC. Au sein des mathématiques, c'est la synergie entre les mathématiques théoriques et les mathématiques appliquées qui est la clé du succès.

5 – LA CRYPTOGRAPHIE CLASSIQUE ET QUANTIQUE

L'adoption récente d'un nouveau standard cryptographique a été l'occasion de mettre en lumière certains aspects de cette technologie. Les algorithmes utilisés se classent en deux catégories : algorithmes à clef secrète et algorithmes à clef publique. Pour les premiers, l'expéditeur et le récipiendaire possèdent la même clef, celle qui permet au premier de crypter le message et au second de le décrypter. Un réseau à grand nombre de participants exige un grand nombre de clefs, ce qui pose des problèmes de stockage. La question cruciale est ici : comment les intervenants se communiquent-ils la clef secrète ? Les algorithmes à clef publique reposent sur des bases mathématiques très sophistiquées, telles que la factorisation de grands entiers, le problème du Log discret et les analogues de ces questions pour les courbes elliptiques (où il est nécessaire de calculer rapidement le nombre de points rationnels d'une courbe elliptique définie sur un corps fini). Ces systèmes sont toutefois soumis à de nombreuses attaques et doivent être constamment améliorés. Parallèlement se développent les techniques de signatures électroniques et le secteur d'activité complètement nouveau qu'est le rôle joué par les autorités de certification dont la tâche est d'authentifier le possédant et les caractéristiques d'une clef publique.

Mais l'enjeu de l'avenir est peut-être la cryptanalyse quantique. Peter Shor, qui fut récipiendaire du Prix Nevanlinna 1998, est un des fondateurs de l'informatique et de la cryptographie quantiques. Il a développé des méthodes basées sur la physique quantique pour factoriser en temps polynomial de grands nombres ou pour résoudre le problème du Log discret même dans le contexte des variétés abéliennes. Le principal obstacle à sa mise en œuvre est l'absence d'ordinateur quantique ! P. Shor estime qu'il faudra au moins 10 ans avant que nous ne disposions d'un premier coprocesseur quantique. Signalons que le département recherche du Pentagone investit annuellement 5 millions de dollars dans ce projet.

Le jour où la technologie existera, toute la cryptographie à clef publique deviendra obsolète. En France des groupes se constituent, des DESS sont créés dans les départements de Mathématiques de certaines Universités, et il semble urgent d'investir dans cette direction où l'Europe a une réelle carte à jouer.

6 – LA MODÉLISATION MATHÉMATIQUE EN SCIENCES DE LA VIE

La modélisation mathématique déterministe en Sciences de la Vie ou Médicales est un outil fort développé dans la culture anglo-saxonne. Elle l'est encore très peu en France, ou elle n'apparaît que de manière sporadique : circulation du sang dans le cœur ; conception de prothèses médicales ou dentaires ; contrôle du métabolisme, croissance de plantes, cellules ou tumeurs ; propagation spatiotemporelle d'épidémies animales ou humaines, transmission inter-espèces ; apparition et persistance de maladies émergentes et leur contrôle, etc.

Les outils mathématiques utilisés sont souvent empruntés aux systèmes dynamiques en temps discret ou continu, couplés avec des méthodes numériques permettant une validation ultérieure. Dans l'état actuel des connaissances , ceux-ci ne sont pas nécessairement toujours très efficaces dans l'optique d'une modélisation de systèmes biologiques complexes présentant un minimum de réalisme. Un gros effort de conceptualisation est ici nécessaire. Ce travail fédère en général des mathématiciens impliqués individuellement (ou de petites équipes) en collaboration pluridisciplinaire autour de projets ciblés financés par des actions spécifiques de type ACI inter EPST.

La génomique est aussi un domaine qui requiert des mathématiques et tout particulièrement des statistiques. Il arrive chaque jour 20 millions de nouvelles lettres dans les banques de données de séquençage, et le génome complet d'un ou deux organismes est publié chaque mois. Analyser cette avalanche de données ne peut se faire que si l'on a mis au point des procédures statistiques capables de repérer l'information – ou une grosse partie de l'information – contenue dans ces séquences. Cette application à la biologie a suscité un certain nombre de travaux théoriques sur les chaînes de Markov cachées. Une fois un segment de génome suspecté d'être un gène, il convient de le comparer aux gènes déjà présents dans les banques de données. Ceci a conduit à la conception de toute une panoplie d'algorithmes, et la significativité statistique des résultats est, encore aujourd'hui, l'objet de recherches mathématiques. D'autres procédures statistiques sont aussi développées en génomique.

Un phénomène vraiment nouveau est l'apparition de mathématiques très théoriques pour comprendre la logique du codage génétique et sa géométrie. Il est bien trop tôt pour se prononcer sur l'efficacité et la pertinence des approches développées, mais cette attitude est révélatrice d'une certaine ouverture dont nous ne pouvons que nous féliciter.

Ce travail mobilise un nombre important d'équipes en France, pour beaucoup structurées par l'Action IMPG (Informatique, Mathématiques, Physique pour la Génomique) du Ministère. Les Génopoles, en particulier celle d'Évry, coordonnent les efforts des nombreux organismes impliqués : CNRS, INRA, INSERM, Institut Pasteur, Universités, etc.

7 – CONCLUSION

Ce bref aperçu de quelques aspects des mathématiques françaises montre que les mathématiciens peuvent intervenir et interviennent dans la société : comme les autres scientifiques, ils produisent de la technologie, et de plus en plus. Les autorités américaines ont compris que les Mathématiques les plus fondamentales peuvent être la clef de nombreuses technologies d'avenir, en particulier – mais pas seulement – celles liées aux STIC. Le CNRS peut compenser l'afflux de dollars qui submergent certains projets par un recrutement important de chercheurs ; et la disponibilité et l'indépendance d'esprit qu'il permet sont des atouts considérables.

Il faut aussi insister sur le fait que les mathématiques fondamentales développées dix ans auparavant pour leur esthétique et leurs aptitudes à résoudre des problèmes mathématiques abstraits produisent des applications inattendues. Il n'en est que plus important de maintenir l'expertise française dans cette discipline.

Épilogue : Au moment où nous terminons ce rapport (novembre-décembre 2002), le premier brevet de cryptologie basé sur une utilisation des modules de Drinfeld vient d'être déposé par une équipe française et accréditer. Une quinzaine d'année seulement a séparé l'introduction de ces objets géométriques hautement abstraits et leur utilisation à des fins industrielles, illustrant parfaitement un des propos de ce rapport.