DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX,
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.

CVM COMMENTARIIS C. G. BACHETI V. C.
& obſeruationibus D. P. de FERMAT *Senatoris Toloſani.*

Acceſſit Doctrinæ Analyticæ inuentum nouum, collectum
ex varijs eiuſdem D. de FERMAT Epiſtolis.

TOLOSÆ,
Excudebat BERNARDVS BOSC, è Regione Collegij Societatis Ieſu.

M. DC LXX.

Autour du théorème de Fermat

Catherine Goldstein

L'annonce en juin 1993 lors d'un colloque à Cambridge qu'Andrew Wiles avait démontré le théorème de Fermat a été abondamment reprise par la presse. Outre les exposés donnés à cette occasion, Wiles a soumis pour publication un épais manuscrit ; les rapporteurs ayant apparemment trouvé au moins une lacune importante, ce texte n'a toujours pas été diffusé plus largement (février 1994). Mais l'intérêt de la démarche elle-même et les résultats qu'elle permet d'ores et déjà d'obtenir justifient certainement qu'on en explique les idées. L'événement constitue en outre une bonne occasion pour faire le point sur l'histoire du théorème (ou plutôt de sa preuve) : en effet, des travaux historiques récents ont permis de mieux comprendre certains aspects des recherches qui s'y connectent.

Rappelons tout d'abord qu'il s'agit de prouver que l'équation $a^n + b^n = c^n$ n'a pas de solutions entières (a, b, c) non nulles, dès que $n \geq 3$. Des extensions à d'autres ensembles de nombres ont été également étudiées (voir [Dickson], chapitre XXVI), mais le théorème serait évidemment faux par exemple pour a, b, c réels : il faut trouver un moyen de se concentrer sur les solutions entières (ou rationnelles), c'est-à-dire faire de l'arithmétique. Par ailleurs, si le théorème est démontré pour un exposant n, il l'est aussitôt pour tous ses multiples, et il suffit donc en principe d'étudier les cas $n = 4$ et $n = p$ premier (quoique d'autres cas, plus faciles, aient pu être étudiés d'abord).

Autour de l'histoire du théorème de Fermat

1. Méthodes

Il est d'usage de distinguer trois grandes familles d'approche [1] :

• les méthodes dites maintenant élémentaires, c'est-à-dire fondées uniquement sur des manipulations algébriques, des congruences et études de divisibilité, la fameuse méthode de descente infinie sur laquelle nous reviendrons. C'est ainsi que les premiers exposants, $n = 4$, 3, 5 et 7 ont été traités.

• la théorie algébrique des nombres. En expansion surtout à partir de la deuxième moitié du 19ème siècle, elle repose sur l'étude des anneaux de nombres formés à partir de racines de certaines équations algébriques, par exemple des anneaux de nombres cyclotomiques, nombres (complexes) de la forme $\sum_{i=0}^{p-1} a_i \zeta^i$ où $a_i \in \mathbf{Z}$ et où ζ est une racine primitive p-ième de l'unité. C'est dans ce cadre qu'ont été traités la plupart des autres cas connus jusqu'à ces dernières années (plusieurs

[1] Ce découpage change souvent quand une nouvelle méthode mathématique s'introduit, rapprochant du même coup d'anciens résultats dans la même catégorie et remodelant rétrospectivement les perspectives ; on pourra s'en convaincre déjà en comparant les tables des matières de [Noguès], [Edwards], [Ribenboim], tous écrits avant que les méthodes géométriques ne commencent à tenir leurs promesses en ce qui concerne le théorème de Fermat.

centaines de milliers d'exposants).

• les méthodes géométriques en arithmétique. Elles se sont surtout développées pendant ce siècle et tout particulièrement depuis les années 60. Elles consistent à interpréter un problème numérique (par exemple celui posé par l'équation de Fermat) en termes géométriques, à l'aide d'une courbe particulière, et à appeler à la rescousse les techniques, souvent très sophistiques, issues de la géométrie algébrique. Deux voies principales peuvent être distinguées pour aborder ainsi le théorème de Fermat : l'une consiste simplement à étudier les (éventuels) points à coordonnées rationnelles de la courbe d'équation $x^p + y^p = 1$ dans le plan (x, y) et a permis à Faltings en 1983 de prouver que l'équation de Fermat n'avait, pour p fixé ≥ 5, au plus qu'un nombre fini de solutions primitives (c'est-à-dire avec a, b, c premiers entre eux). L'autre voie conduit à essayer de montrer que si $a^p + b^p = c^p$ était un contre-exemple au théorème de Fermat, les propriétés de la courbe définie par l'équation $y^2 = x(x - a^p)(x + b^p)$ contrediraient ce qu'on attend de courbes cubiques de ce type. C'est dans cette deuxième voie que se situe le travail de Wiles, et nous l'examinerons plus en détail dans la deuxième partie de ce texte.

Parce que cette variation des méthodes correspond à peu près à l'ordre chronologique, il serait tentant de les voir dans une stricte succession, chacune rendant obsolète la précédente. La réalité est plus compliquée et un coup d'œil aux *Mathematical Reviews* montre qu'aucune des approches n'a été complètement abandonnée. Mais elles définissent plutôt des groupes différents, aux intérêts et aux objectifs multiples, et si des résultats intéressants ont été atteints pendant ce siècle par des méthodes élémentaires (voir [Ribenboim]), obtenir ainsi la première preuve *générale* paraît tout à fait utopiste, nous y reviendrons. La théorie algébrique des nombres reste quant à elle, et indépendamment du théorème de Fermat, un domaine très important de recherches. Les deux approches géométriques se sont développées presque simultanément et rien ne dit que c'est la dernière proposée qui founira la première une preuve de l'énoncé complet. Les interactions sont d'ailleurs nombreuses et parfois inattendues.

De plus, d'autres branches importantes de la théorie des nombres contemporaine ont fourni des renseignements sur le théorème de Fermat. En 1985, Adleman et Heath-Brown ont prouvé, à partir d'un critère qui généralise celui de Sophie Germain, qu'une minoration convenable du nombre des premiers $l \equiv 2 \bmod 3$, $l \leq x$, tels que $l-1$ ait un grand facteur premier (en l'occurrence un facteur premier supérieur à x^δ avec $\delta > 2/3$) entraînerait l'existence d'une infinité d'exposants premiers p pour lesquels l'équation de Fermat n'admettrait pas de solutions non nulles premières[2] à p. L'estimation était connue pour certains $\delta < 2/3$, par exemple $\delta < 0,65635$ (Iwaniec-Deshouillers), mais la barre de $2/3$, nécessaire pour l'application au théorème de Fermat, fut franchie par Fouvry en 1985, à l'aide de techniques analytiques très subtiles, voir [Deshouillers].

Une autre tentation, accentuée par la fascination qu'a exercée le théorème de Fermat sur beaucoup de non-spécialistes (et sur beaucoup de mathématiciens et mathématiciennes également !), est

[2] Depuis les travaux de Sophie Germain, au début du 19ème siècle, on distingue souvent deux cas dans la démonstration du théorème de Fermat : le premier cas consiste à montrer qu'il n'y a pas de solutions premières à l'exposant, le second cas qu'il n'y en a pas où l'exposant et l'un des nombres a, b, c ait un diviseur commun. Germain elle-même avait prouvé le premier cas pour tous les nombres premiers jusqu'à 100, par des méthodes élémentaires et astucieuses. Mais le second cas pour les mêmes premiers a exigé les méthodes de Kummer qui relèvent de la théorie algébrique des nombres. Par ailleurs, Terjanian a prouvé que le premier cas était vrai pour tous les exposants *pairs* en 1977, par des méthodes élémentaires; par contre, le premier cas pour une infinité d'exposants *premiers* a exigé le travail de Fouvry évoqué dans le texte. Comme on le constate, la difficulté relative des différents résultats et leur accessibilité par une méthode ou une autre ne va pas de soi. On trouvera de nombreux autres exemples dans [Ribenboim] et dans la suite de ce texte.

d'interpréter tous ces développements comme la conséquence d'un intérêt permanent et exclusif pour ce théorème depuis le 17ème siècle. C'est plutôt l'inverse qui est vrai : la célébrité de l'énoncé, surtout à partir du 19ème siècle (où des Prix publics ont été offerts pour sa solution), a engagé plus d'une personne à y essayer ses forces ou ses nouvelles méthodes. Comme on le verra abondamment, d'autres questions et d'autres motivations sont à l'arrière-fond de beaucoup de ces approches, et les avances faites sur le théorème de Fermat ne sont qu'une conséquence glorieuse, mais surtout folklorique. Un des problèmes auxquels est confronté le ou la spécialiste est le manque de prise sur un énoncé qui semble pouvoir être accommodé à tant de sauces différentes : il est agréable (et motivant) de trouver des cadres dans lequel le théorème de Fermat se relie à des phénomènes plus généraux, et de préférence plus techniques, et d'y emprunter des outils de travail adéquats.

Il n'est évidemment pas question d'aborder tous ces aspects ici et je me contenterai de résumer quelques résultats mathématiques et historiques récents et bien documentés (à ma connaissance) en renvoyant pour des compléments aux livres de [Edwards] et [Ribenboim].

2. Le cas Fermat

Une question inévitable quand on évoque ce théorème est : "Fermat avait-il démontré son théorème?" Les réponses ont varié et les reconstitutions les plus fantaisistes, de la "preuve" de Fermat ou même de ses "erreurs" [3] ont été proposées. La conviction que cette introuvable démonstration a bel et bien existé (et donc que le résultat est accessible par les méthodes, considérées maintenant comme élémentaires, que pouvait posséder Fermat) continue à animer les efforts des amateurs qui inondent depuis près d'un siècle et demi les Académies des Sciences et autres institutions mathématiques de tentatives de preuves. La conclusion des historiens qui ont étudié en détail les papiers de Fermat (Paul Tannery, co-responsable de la la publication des *Œuvres* de Fermat, Jean Itard plus récemment, voir [Itard]) est que le mathématicien toulousain n'avait sans doute pas de preuve complète pour tous les exposants et qu'il le savait probablement lui-même. Voici l'état du dossier.

Le seul témoignage fourni par Pierre de Fermat de l'existence possible d'une telle preuve est la fameuse note écrite en marge de son exemplaire des *Arithmétiques* de Diophante, dans l'édition traduite en latin et annotée de Claude Gaspard Bachet de Méziriac de 1621. L'exemplaire même n'a jamais été retrouvé, mais Samuel de Fermat, le fils du mathématicien, a publié en 1670, cinq ans après la mort de son père, une nouvelle édition des *Arithmétiques*, incluant les anciens commentaires de Bachet et les nouveaux, faits par son père. La note qui nous concerne suit un commentaire de Bachet sur la fabrication de triplets de nombres vérifiant la relation $a^2 + b^2 = c^2$: "Cubum autem in duos cubos", écrit Fermat, "aut quadratoquadratum in duos quadratoquadratos et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere : cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet" [4] . Voilà, je le

[3] La légende selon laquelle Fermat aurait par exemple utilisé une factorisation de l'équation $a^n + b^n = c^n$ en nombres complexes, à la manière de Kummer, sans se rendre compte que l'unicité de la factorisation en nombres irréductibles ne valait plus dans ce cas, n'a aucun fondement. On trouve bien des nombres imaginaires utilisés pour résoudre des équation algébriques dans des travaux antérieurs au $17^{ème}$ siècle, mais dans un contexte tout différent. Il n'y a aucune trace de leur emploi chez Fermat. Soulignons au passage que l'historien Edwards qui a étudié en grand détail les papiers de Kummer a aussi mis en doute l'"erreur" analogue qu'on attribue souvent à Kummer, à cause d'un récit postérieur de Hensel, voir [Edwards, 1975 & 1977].

[4] "Mais diviser un cube en deux cubes, une puissance quatrième en deux puissances quatrièmes, et généralement une puissance quelconque au-delà de la quatrième à l'infini en deux de même exposant est impossible : j'ai trouvé une preuve vraiment merveilleuse de cela. L'étroitesse de la marge ne la contiendrait pas."

répète, l'unique mention qui puisse faire croire à l'existence d'une preuve générale.

On trouve par contre dans la correspondance de Fermat plusieurs références (avec l'affirmation qu'il en possède une preuve) aux équations $a^3 + b^3 = c^3$ et $a^4 + b^4 = c^4$, soit ensemble, soit séparément, éventuellement avec d'autres problèmes connexes (par exemple le fait qu'un triangle rectangle dont les côtés sont mesurés par des nombres entiers, comme 3, 4, 5, ne peut avoir pour surface le carré d'un entier)[5] . Enfin en 1659, Fermat a écrit un bilan de ses recherches sur les nombres à son ami Carcavi : il ne fait figurer que le cas des cubes (et le problème sur l'aire des triangles qui peut être relié au cas de l'exposant 4) ; n'y figure pas le cas général, pas même à titre de conjecture ou d'espoir futur.

Deux descriptions des événements sont donc possibles : l'une consisterait à croire que les marges du Diophante ont été annotées tardivement, après la lettre à Carcavi, par exemple en guise d'aide-mémoire en vue d'une publication — c'est assez peu vraisemblable, à cause de leur formulation et des renseignements que nous avons par ailleurs sur certaines d'entre elles via la correspondance. L'autre possibilité[6] est que les notes du Diophante aient été écrites au moment de l'étude du livre par Fermat autour des années 1640, comme le suggère la correspondance. Dans ce cas, Fermat a peut-être cru un peu rapidement être en possession d'une preuve générale, à partir d'une preuve pour les premiers exposants[7] , mais il semble s'être rendu compte ensuite des difficultés. Un dernier argument, que je n'ai jamais vu mentionné mais qui est probablement bien connu, est que les preuves pour les exposants 3 et 4 reposent en dernière instance sur une étude de la divisibilité des nombres de la forme a^2+b^2, a^2+2b^2 ou a^2+3b^2. Cette étude devient beaucoup plus compliquée pour a^2+5b^2, qui, dans l'optique d'une généralisation de la méthode, est le cas suivant à considérer ; on sait que Fermat s'est beaucoup intéressé à ces questions dans la suite de sa carrière, dans les années 50 en particulier. Ceci pourrait expliquer ses espoirs, puis son silence sur les exposants supérieurs à 4. Comme toujours en histoire, seul un nouveau document (établissant par exemple la datation des notes du Diophante) pourrait confirmer ou infirmer définitivement cette conclusion.

3. Les premiers exposants

Nous n'avons que très peu d'indications même dans les premiers cas sur les preuves de Fermat. La lettre-bilan de 1659 affirme qu'ils relèvent de la méthode de descente infinie et nous possédons par ailleurs une démonstration rapidement esquissée du théorème sur l'aire des triangles, dans une autre note du Diophante. Ce dernier résultat est lié comme je l'ai déjà dit au théorème de Fermat pour l'exposant 4, plus précisément, au fait qu'une différence de puissances quatrièmes ne peut être un carré (et donc a fortiori une puissance quatrième)[8] . Voici comment (les notations algébriques

---

[5] Dans [Fermat, OC], t. II, lettres XII à Mersenne, XXXIX au même, XCI à Digby, CI à Carcavi, entre autres. Notons que la première lettre a d'abord été datée de 1637 par Charles Henry, co-éditeur des *Œuvres* de Fermat, puis de 1636 dans la publication même de ces oeuvres, puis de 1638 par Itard et indépendamment Cornélis de Waard, l'éditeur de la correspondance de Mersenne. La première date est reproduite dans [Dickson] et, sans doute d'après lui, dans de nombreux écrits sur le théorème de Fermat. Indépendamment des fluctuations mentionnées entre 1636 et 1638, il ne s'agit que de la première mention écrite que nous ayons des *premiers* cas du théorème, sans allusion aucune au cas général.

[6] Ceux qui plaident en faveur de l'existence d'une preuve générale possédée effectivement par Fermat font l'économie de cette alternative ! Ils voient l'œuvre de Fermat comme un tout et ne s'appuient pas en général sur une analyse de l'évolution de Fermat, qui est pourtant très perceptible, comme l'a montré [Itard], voir aussi [Weil].

[7] Ceci concorderait d'ailleurs avec la manière dont les problèmes arithmétiques étaient traités avant Fermat, et en particulier dans le corpus euclidien.

[8] On pourrait prouver alternativement qu'une somme de puissances quatrièmes ne peut être un carré, ce résultat figure sans indication de preuve dans une autre note, voir par exemple [Edwards].

sont modernes et ne correspondent pas du tout à l'écriture de Fermat) :

Un triangle rectangle en nombres est simplement un triplet de nombres entiers vérifiant $r^2 + s^2 = t^2$. On peut supposer que ces trois nombres sont premiers entre eux, que r et t sont impairs et s est pair. Il n'est pas très difficile de montrer (par des arguments de parité et de factorisation par exemple) qu'il existe deux entiers p et q premiers entre eux et de parité différente tels que

$$r = p^2 - q^2, \quad s = 2pq, \quad t = p^2 + q^2.$$

L'aire du triangle rectangle est alors $pq(p^2 - q^2)$: or, p, q, $p^2 - q^2$ sont premiers entre eux et si l'aire, c'est-à-dire leur produit, était un carré, chacun de ces nombres devrait être un carré, soit

$$p = u^2, \quad q = v^2, \quad p^2 - q^2 = w^2,$$

donc $u^4 - v^4 = w^2$. Si cette dernière équation n'a pas de solutions, l'équation $a^4 + b^4 = c^4$ n'en a pas non plus [9] .

La preuve que l'aire de ces triangles rectangles n'est pas carrée (et donc le cas de l'exposant 4 du théorème de Fermat), esquissée en note par Fermat, repose sur des factorisations, l'étude des nombres de la forme $p^2 + 2q^2$ et $p^2 + q^2$, et la descente infinie : Fermat montre en effet que s'il existait un tel triangle (ou une solution à l'équation $u^4 - v^4 = w^2$), il en existerait un(e) autre, en entiers également, mais strictement plus petits. Puisqu'une suite d'entiers ne peut décroître indéfiniment, il aboutit à une contradiction [10] .

Le cas de l'exposant 3 n'est jamais explicité dans l'œuvre de Fermat; on la trouve chez Euler avec des principes très proches, voir par exemple [Weil].

Remarques

J'ai argumenté ailleurs, [Goldstein, 1989 & 1993], que Fermat se trouve à la jonction de plusieurs attitudes face à l'arithmétique, ce qui le singularise parmi ses contemporains (au moins ceux dont nous avons des traces) : il est un des rares à essayer de développer les potentialités que l'algèbre offre dans ce domaine précis. Il est peut-être nécessaire de rappeler ici que c'est seulement au cours du 17ème siècle que l'algèbre est vraiment assimilée par les mathématiciens occidentaux, assimilation à la fois technique et culturelle dans laquelle les *Arithmétiques* de Diophante jouent un rôle spécial, voir [Cifoletti] et [Morse]. Mais Fermat est aussi conscient des exigences particulières que pose l'arithmétique (pour lui plus précisément le travail sur les entiers) et la descente infinie lui semble un moyen de concilier ses objectifs [11] .

Cette double approche n'a pas fait vraiment école avant le 18ème siècle, bien que des traces d'intérêt continu pour l'un ou l'autre aspect soient tout à fait claires avant (voir [Goldstein, 1993]). L'histoire du théorème de Fermat au 18ème siècle est à la fois très connue et peu documentée : à

---

[9] Le fait que u, v, w proviennent de la situation des triangles rectangles impose en principe certaines conditions de parité et ne fournissent pas le cas général de l'équation $u^4 - v^4 = w^2$. Mais il est facile de l'en déduire, voir [Goldstein, 1994b].

[10] Les détails figurent dans presque tous les livres sur la théorie des nombres et son histoire, par exemple [Edwards], [Ribenboim], [Dickson], [Weil]; voir aussi [Goldstein, 1994b] pour une étude comparative des différentes présentations.

[11] Cette interprétation confirme d'ailleurs l'analyse faite par Rashed sur la théorie des nombres en langue arabe, où une telle confluence ne semble pas avoir eu lieu : des mathématiciens intéressés par la théorie des nombres à la manière euclidienne ont exploré l'impossibilité qu'un cube entier soit somme de deux cubes avant le 11ème siècle, mais semblent distincts de ceux qui se sont intéressés à l'analyse diophantienne et à la recherche des solutions d'équations algébriques, voir [Rashed], chapitre IV.

part les recherches d'Euler, de Lagrange, de Legendre, qui ont conduit entre autres aux preuves pour les exposants suivants, l'étendue du travail accompli par d'autres mathématiciens, la place de ces recherches dans les mathématiques de l'époque, les options fondamentales qu'elles recouvrent, n'ont guère été étudiées. Un travail très important a par contre été effectué pour la période suivante (1840-1880 disons) autour des résultats et des méthodes de Kummer, puis de ses successeurs, en particulier par Harold Edwards [12] .

4. La théorie algébrique des nombres

Je me contenterai, en m'appuyant en partie sur les conclusions d'Edwards, d'expliquer la place de Kummer et de ses successeurs dans l'histoire du théorème de Fermat, mais aussi quelques-unes des idées mathématiques importantes liées à ce courant (développées ensuite sans lien direct avec la théorème de Fermat, mais qui ont alimenté certaines recherches utilisées par Wiles...).

Tout d'abord, si Kummer a écrit, alors qu'il était enseignant dans un lycée, un petit article sur le théorème de Fermat [13] , son intérêt principal vers 1844 semblait s'être tourné vers les lois de réciprocité, dans la tradition inaugurée par Gauss et Jacobi, voir [Edwards, 1975 & 1977], et [Neumann].

La plus simple de ces lois est la loi de réciprocité quadratique : si p et q sont deux nombres premiers distincts impairs, alors p est congru à un carré modulo q si et seulement si q est congru à un carré modulo p, sauf si p et q sont tous deux de la forme $4n+3$, auquel cas les deux possibilités s'excluent mutuellement et alors $p \equiv s^2 \,\text{modulo}\, q$ si et seulement si $q \equiv -s'^2 \,\text{modulo}\, p$. Cette loi apparaît naturellement lorsqu'on s'intéresse à la représentation des nombres comme somme de carrés et de multiples de carrés (plus généralement de formes quadratiques), problème qui s'inscrit, comme je l'ai dit plus haut, dans une tradition où figurent Fermat, Euler, Lagrange [14] . Par exemple, si un nombre premier p s'écrit sous la forme $s^2 + nt^2$ (avec s et t entiers non nuls), il doit être congru à un carré modulo tous les diviseurs de l'entier n. La loi de réciprocité intervient efficacement lorsqu'on veut déterminer pour quels n un nombre premier p fixé peut se mettre sous la forme indiquée.

Gauss lui-même avait remarqué vers 1830 que le recours aux nombres $\{a + bi \mid a \in \mathbf{Z},\, b \in \mathbf{Z}\}$ donne une formulation plus élégante de la loi de réciprocité et permet de démontrer une extension à une loi biquadratique du même type. Il dut pour cela prouver que les nombres $\{a + bi\}$ vérifient des propriétés analogues aux entiers usuels, en particulier se décomposent de manière unique en produit de nombres irréductibles, à une unité près (les unités sont ici ± 1, $\pm i$). La généralisation à des lois supérieures imposait d'étudier de la même façon les anneaux de nombres que nous appelons maintenant cyclotomiques, c'est-à-dire de la forme $\{\sum a_i \zeta^i \mid a_i \in \mathbf{Z},\, \zeta^p = 1\}$. Pour de tels entiers, la factorisation en éléments irréductibles n'est plus unique mais, comme il est bien connu, Kummer eut l'idée de récupérer cette unicité en introduisant de nouveaux nombres, "idéaux".

L'analogie la plus frappante de cette construction a été fournie par Kummer lui-même : c'est

---

[12] Voir [Edwards, 1975 & 1977]. On y trouvera aussi des informations sur la situation en France, en particulier sur les recherches de Cauchy et Lamé et le prix de 1850 à l'Académie des Sciences. Pour le premier prix proposé en 1816 et les résultats de Sophie Germain, voir [Bucciarelli & Dwork].

[13] Sur le cas des exposants pairs. Comme nous l'avons déjà dit, même si le cas des exposants premiers est suffisant, il y avait déjà quelques raisons à l'époque de croire les exposants pairs plus accessibles. Lejeune-Dirichlet avait ainsi montré que le théorème de Fermat était vrai pour l'exposant 14 avant qu'on sache le faire pour 7.

[14] Voir [Weil] et aussi [Davenport] pour une présentation limpide de cette loi et de quelques-unes de ses preuves; on en connaît plus d'une centaine maintenant ! Pour des applications technologiques de cette loi, voir [Schroeder].

celle d'une réaction chimique où plusieurs équilibres avec des corps différents semblent pouvoir se produire. Une analyse plus fine montre que les mêmes éléments, dans les mêmes proportions, se trouvent présents sous les combinaisons variées qui forment les corps, les seuls observables en première instance. De même l'analyse en nombres idéaux permet de récupérer à leur niveau une unicité de décomposition qui n'est pas observable sur les nombres de départ. Ce phénomène a donné lieu à plusieurs réinterprétations [15] : celle de Kronecker est fondamentale pour les liens avec la géométrie algébrique, mais celle de Dedekind est peut-être plus familière. Un idéal I d'un anneau A unitaire est simplement un A-module inclus dans A, autrement dit, il est stable par addition et par multiplication par n'importe quel élément de l'anneau A ; l'ensemble des multiples dans A d'un élément a est un idéal, appelé idéal principal. Le produit de deux idéaux est l'idéal engendré par les produits d'un élément du premier par un élément du second (en particulier, il contient les sommes de tels produits et est inclus dans l'intersection des deux idéaux). Un idéal est dit premier si $\frac{A}{I}$ est intègre (autrement dit, si x ou $y \in I$ dès que $xy \in I$.). Dans un anneau comme celui des entiers cyclotomiques, $\mathbf{Z}[\zeta_p]$, où ζ_p est une racine p-ième de l'unité, tout idéal se décompose de manière unique en un produit fini d'idéaux premiers. De plus, une puissance convenable d'un idéal (le nombre de classes) est un idéal principal [16] .

L'application au théorème de Fermat vient de la factorisation :

$$c^p = a^p + b^p = (a+b)(a+b\zeta)\cdots(a+b\zeta^{p-1}),$$

où l'on a posé $\zeta = \zeta_p$. On peut voir cette factorisation comme une factorisation en idéaux principaux et le produit des idéaux $(a + b\zeta^i)$ est donc une puissance p-ième $(c)^p$. On aimerait en déduire que chaque facteur $(a + b\zeta^i)$ est (à des diviseurs communs contrôlables près) une puissance p-ième, mais ce n'est alors que la puissance p-ième d'un idéal (pas nécessairement principal). En étudiant la structure de l'anneau des entiers cyclotomiques $\mathbf{Z}[\zeta]$ (ses classes d'idéaux et ses unités en particulier), Kummer fut capable de prouver que si p est premier au nombre de classes de $\mathbf{Z}[\zeta]$ (un tel nombre premier est dit *régulier*), le théorème de Fermat est vrai pour l'exposant p. Il avait l'espoir qu'une infinité de p étaient réguliers, mais on ne connaît toujours pas la réponse à cette question : par contre, on sait depuis le travail de Jensen en 1915 qu'il existe une infinité de nombres premiers irréguliers [17] . Kummer lui-même donna un critère très important pour calculer le nombre de classes et montra le théorème de Fermat pour tous les nombres premiers inférieurs à 100 (sauf trois exceptions irrégulières 37, 59 et 67). C'est pour ce résultat que Kummer obtint en 1857 le

---

[15] Voir [Edwards, 1992] pour une étude de leur variété et les motivations diverses de ces réinterprétations.

[16] On peut en fait définir un idéal fractionnaire du corps des fractions K d'un anneau A intègre : c'est un A-module non nul de K tel qu'il existe un élément d de A tel que dI est inclus dans A. L'ensemble des idéaux fractionnaires forme alors un groupe pour le produit des idéaux. Un idéal fractionnaire principal est de la forme aA, avec $a \in K$. Le quotient du groupe des idéaux fractionnaires par les idéaux principaux, le "groupe des classes" d'idéaux, dans un corps comme $\mathbf{Q}(\zeta_p)$, est fini et son nombre est le "nombre de classes" évoqué dans le texte. Tout idéal de $\mathbf{Z}$ étant principal, le nombre de classes de $\mathbf{Q}$ est 1 ; il en est de même par exemple pour $\mathbf{Q}(\zeta_3)$ ou $\mathbf{Q}(\zeta_{19})$, mais le nombre de classes de $\mathbf{Q}(\zeta_{29})$ est 8, voir [Washington].

[17] Voir aussi [Washington], p. 62. Ceci ruine donc l'espoir de prouver le théorème de Fermat en appliquant directement le résultat de Kummer et en étudiant un nombre fini d'exceptions connues. Pour un nombre premier irrégulier donné *explicitement*, on peut en effet étudier la structure de l'anneau cyclotomique correspondant et conclure. On dispose aussi de nombreux raccourcis élaborés par les successeurs de Kummer, dont certains constituent des avancées théoriques importantes, voir [Ribenboim].

Prix de l'Académie des Sciences. Soit B_n le n-ième coefficient de la série

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}.$$

Tous les B_n pour n impair ≥ 3 sont nuls. Les autres coefficients se calculent facilement par récurrence. Kummer montra qu'un nombre premier est régulier si et seulement s'il ne divise aucun des numérateurs des coefficients B_{2k}, avec $k = 1, \cdots, \frac{p-3}{2}$. Par exemple, 37 divise le numérateur de B_{32}, qui a 42 chiffres et que je ne retranscris donc pas ici .... Ce résultat a bien sûr stimulé beaucoup de recherches et de calculs sur les nombres B_n (les "nombres de Bernoulli") [18] .

Les développements de ces idées sont multiples : la notion d'idéal par exemple a bien sûr débordé du cadre strict de la théorie des nombres. L'étude des lois de réciprocité a donné lieu entre autres à la théorie du corps de classes. Soit un anneau A, extension convenable de $\mathbf{Z}$, par exemple l'anneau des entiers cyclotomiques, et (p) l'idéal principal de A composé des multiples dans A d'un nombre premier p usuel. On se propose d'étudier comment l'idéal (p) se décompose en idéaux premiers de A et jusqu'à quel point ces décompositions caractérisent l'extension A. Par exemple, l'idéal engendré par p, un nombre premier impair différent de 7, dans l'anneau $\{a + b\sqrt{7} \mid a \in \mathbf{Z}, b \in \mathbf{Z}\}$ se décompose en deux idéaux premiers distincts exactement si 7 est un carré, r^2, modulo p : dans ce cas, $(7) = \wp_- \times \wp_+$, où $\wp_\mp$ est l'idéal engendré par p et $\sqrt{7} \mp r$. La loi de réciprocité donne une forme plus commode à cette condition, puisqu'elle permet d'étudier une fois pour toutes les carrés modulo 7 au lieu d'avoir à vérifier si 7 est un carré modulo différents nombres premiers p. La théorie du corps de classes réussit à étendre ce type de résultats à d'autres extensions de $\mathbf{Z}$, qui sont engendrées par des racines d'équations algébriques dont le groupe de Galois (voir plus loin) est commutatif, et à décrire ces extensions à partir de la décomposition de certains idéaux (p), p un nombre premier usuel [19] . Un programme, dit de Langlands, entamé dans les années 60 et loin d'être complètement réalisé à l'heure actuelle, fournit désormais un cadre pour étudier le cas général, voir [Gelbart].

Un autre aspect très important des travaux de Kummer est la relation entre nombres de Bernoulli et nombre de classes. Les nombres de Bernoulli sont liés à des valeurs spéciales de la célèbre fonction ζ de Riemann par la formule suivante, valable pour tout entier $k \geq 1$:

$$B_{2k} = (-1)^{k-1} \frac{2(2k)!}{(2\pi)^{2k}} \zeta(2k).$$

D'autres fonctions généralisant la fonction ζ, les fonctions L dont nous verrons des exemples dans le cadre des recherches de Wiles, prennent (à des facteurs bien contrôlés près, comme ici les puissances de 2π) des valeurs rationnelles comme les nombres de Bernoulli ; l'interprétation de ces valeurs en fonction de structures algébriques, comme le groupe de classes, a donné lieu à quelques-unes des plus importantes conjectures actuelles en théorie des nombres, comme celles de Birch et Swinnerton-Dyer, de Beilinson, de Bloch-Kato.

---

[18] Une présentation moderne et très claire de toutes ces questions figure dans [Washington].

[19] Sur l'histoire de la loi de réciprocité et de la théorie du corps de classes, voir les travaux en cours de Frei, dont [Frei].

5. Intermède

Avant de nous tourner vers des approches plus contemporaines, je voudrais souligner ici à quel point la renommée du théorème de Fermat a cru au cours du 19ème siècle, période qui coïncide aussi avec le développement de son historiographie, savante et populaire. Se sont conjugués pour cela l'implantation massive de la théorie des nombres algébriques dans les universités allemandes, surtout Berlin et Göttingen, de grande renommée internationale [20] , les prix prestigieux offerts pour la résolution du problème, mais aussi l'expansion des manuels scolaires avancés et des journaux de niveau intermédiaire, destinés aux élèves des écoles d'ingénieurs ou assimilés et à leurs enseignants. Ceux-ci se sont montrés particulièrement friands de problèmes de géométrie (euclidienne surtout) et d'arithmétique. L'influence réciproque de ces divers facteurs sur la destinée du théorème après les années 1870 n'est pas encore élucidée.

6. Les courbes algébriques, de Mordell à Faltings (et au-delà).

Une interaction de plus en plus forte [21] lie maintenant certaines branches de la théorie des nombres et la géométrie des courbes algébriques (ou des surfaces, des variétés de dimension supérieure et de leurs généralisations). Les problèmes classiques de l'analyse diophantienne (chercher des solutions rationnelles à des équations ou à des systèmes d'équations polynomiales en un nombre fini de variables) peuvent par exemple être traduits en termes géométriques : on cherche alors les points à coordonnées rationnelles sur les courbes (surfaces, etc.) définies dans un espace affine ou projectif adéquat par les équations ou systèmes d'équations en question.

Le degré et le nombre de variables ne sont pas de très bons invariants de ce point de vue : on a envie d'identifier les équations définissant la même courbe et même d'identifier deux courbes birationnellement équivalentes, c'est-à-dire qui diffèrent l'une de l'autre par un changement de coordonnées défini presque partout par des fractions rationnelles. Un exemple standard est celui du cercle d'équation $x^2 + y^2 = 1$ dont un paramétrage classique est

$$x = \frac{2t}{t^2+1}, y = \frac{1-t^2}{1+t^2}.$$

Réciproquement

$$t = \frac{x}{y+1},$$

dès que $y \neq 1$. Il y a donc équivalence birationnelle entre le cercle (défini par une équation à deux variables de degré 2) et la droite (repérée par la coordonnée t, ou par l'équation qu'on voudra, disons $y = 3$ dans le plan x, y), la correspondance envoyant le point d'abscisse t sur le point $x = \frac{2t}{t^2+1}, y = \frac{1-t^2}{1+t^2}$ et réciproquement.

Il est donc d'usage d'introduire un nombre invariant par ces transformations, le genre, qui dépend du degré, mais aussi des éventuels points singuliers de la courbe (un point singulier est un point où la tangente n'est pas définie) : par exemple, la courbe $y^2 = x^3$ et la courbe $y^2 = x^3 + 1$ ont même degré, mais la première est birationnellement équivalente à une droite (en posant $t = y/x$ et réciproquement $x = t^2, y = t^3$), alors que la seconde ne l'est pas, voir [Reid]. Cette différence est

[20] Ce qui ne signifie pas que les idées de Kummer eurent partout le même impact, voir [Edwards, 1975] pour les rapports de Kummer avec les membres de l'Académie des Sciences de Paris et aussi [Goldstein, 1994a] pour les réticences des théoriciens des nombres français aux méthodes des nombres idéaux.

[21] On trouverait certainement des "précurseurs" au siècle dernier, comme Kronecker ou Hilbert et Hurwitz. Poincaré fut l'auteur en 1901 d'un article-programme très important sur ces questions. Un des acteurs essentiels dans notre siècle est André Weil, voir [Weil, OS], tome 1, commentaires p. 521, 542-543 et la lettre à Simone Weil de 1940.

liée à l'existence d'un point singulier sur la première courbe (le point de rebroussement à l'origine) : la paramétrisation s'obtient en effet en intersectant avec la courbe une droite de pente variable t passant par le point singulier. Le genre de la première courbe est 0 (comme celui du cercle, de la droite, des coniques), celui de la seconde 1. Le genre d'une courbe plane sans singularités définie par une équation de degré n est $(n-1)(n-2)/2$, le genre baisse s'il y a des singularités[22] .

En particulier, la courbe $x^n + y^n = 1$ a un genre ≥ 2 dès que $n \geq 4$. Or Mordell avait conjecturé dès 1922 qu'une courbe de genre supérieur ou égal à 2 ne pouvait avoir au plus qu'un nombre fini[23] de points à coordonnées rationnelles. Ce résultat (pour toutes les courbes de genre au moins 2) n'a été démontré qu'en 1982-1983 par G. Faltings, en utilisant tout l'arsenal des méthodes de géométrie algébrique développées depuis les années 60, sous l'impulsion en particulier de Grothendieck. On trouvera dans [Cornell & Silverman] les informations sur cette preuve (à un niveau "recherche spécialisée"). Appliqué au cas particulier où la courbe est $x^n + y^n = 1$, le résultat de Faltings implique donc que l'équation de Fermat n'a au plus qu'un nombre fini de solutions primitives.

La première démonstration de Faltings n'était pas du tout effective, autrement dit ne fournissait pas de méthodes pour trouver concrètement les solutions. Il faut pour cela non seulement une borne du nombre de ces solutions, mais aussi une borne de leur taille éventuelle : il est impossible de chercher un nombre rationnel (en particulier par ordinateur) sans savoir a priori la grandeur de son dénominateur. De nouvelles idées de Bombieri, Vojta, Faltings lui-même, ont permis une preuve alternative, plus simple et plus générale, vers 1991, mais on ne sait toujours pas borner la taille des solutions (même dans le seul cas des équations de Fermat). Ces recherches mobilisent de nombreux spécialistes.

Une autre voie géométrique très différente a été également suivie. Il ne s'agit plus d'interpréter l'équation de Fermat comme équation d'une courbe (de genre supérieur ou égal à 2), mais de déduire d'éventuelles solutions (non nulles) à cette équation les coefficients d'une courbe de genre 1 et de montrer que cette courbe aurait des propriétés invraisemblables. Plusieurs variantes de cette approche ont été essayées, voir [Œsterlé], nous ne parlerons ici que de celle qui débouche sur les récents travaux d'Andrew Wiles. La suite de cet article lui est consacrée.

Autour de la conjecture de Taniyama-Shimura-Weil : l'approche de Wiles.

1. Introduction

J'aimerais ici essayer d'expliquer quelques-unes des idées en cause : le problème principal est que les nombreux objets mathématiques nécessaires (nombres l-adiques, surfaces de Riemann et courbes algébriques, représentations galoisiennes, formes modulaires) sont bien connus des spécialistes, mais ne font pas du tout partie de la formation commune en mathématiques. Il est indispensable,

[22] Ceux et celles qui seraient plus familiers avec les surfaces de Riemann préféreront la définition suivante : une courbe algébrique définit une surface de Riemann, qui est topologiquement un tore à g trous et le genre de la courbe est justement le nombre g de trous. Il existe de nombreuses définitions équivalentes du genre, relevant de domaines très différents des mathématiques : géométrie algébrique, géométrie différentielle, topologie, etc., voir par exemple [Reyssat].

[23] Le cas du cercle convaincra aisément qu'il n'en est pas de même en genre 0 : dès qu'il y a un point rationnel, il y en a une infinité. La situation est plus compliquée en genre 1 : il peut n'y avoir aucun point rationnel, un nombre fini, une infinité. L'approche suivie par Wiles fait appel à une étude des courbes de genre 1, mais pas spécialement de leurs points rationnels. Pour ces questions, voir [Silverman] et aussi [Koblitz].

pour formuler les résultats, d'introduire plusieurs de ces notions : j'ai essayé de le faire dans des paragraphes relativement autonomes en espérant que cela facilitera le lecture. Je n'ai pas réussi à être à la fois exacte et compréhensible partout, et espère seulement que l'un des deux au moins sera toujours vrai.

Le point de départ de Wiles est un résultat spectaculaire de Ken Ribet de 1987, voir [Ribet, 1990 a & b] : celui-ci a prouvé que le théorème de Fermat résulterait d'une conjecture importante en théorie des nombres, la conjecture que nous appellerons ici STW [24] . Cette conjecture concerne les courbes elliptiques définies sur $\mathbf{Q}$, c'est-à-dire les courbes E de genre 1 dont l'équation peut par un choix convenable de coordonnées être écrite sous la forme :

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

où les a_i sont des entiers. Une telle équation définit bien une courbe de genre 1 si elle n'a pas de points singuliers (revoir la formule du genre donnée plus haut!), autrement dit si les dérivées partielles par rapport à x et y de l'équation ne s'annulent pas en même temps. Ceci se traduit facilement par le fait qu'un certain polynôme Δ en les coefficients a_i ne s'annule pas. Par exemple, si $a_1 = a_2 = a_3 = 0$, $\Delta = -16(27a_6^2 + 4a_4^3)$ (on reconnait à une constante près le discriminant usuel d'un polynôme cubique) [25] .

Voici quelques exemples de courbes elliptiques définies sur $\mathbf{Q}$:

$y^2 = x^3 - x$	$\Delta = 2^6$
$y^2 + y = x^3 - x^2$	$\Delta = 11$
$y^2 + y = x^3 - x$	$\Delta = 37$
$y^2 = x(x - a^p)(x + b^p)$, avec $a, b, c \in \mathbf{Z}$, $a^p + b^p = c^p$,	$\Delta = \frac{(abc)^{2p}}{256}$

Le dernier exemple est bien sûr celui qui nous intéressera particulièrement : on constate alors que le discriminant Δ est non nul exactement lorsque a, b, c ne le sont pas non plus, autrement dit, s'ils constituent un contre-exemple au théorème de Fermat.

Ces dernières courbes ont été étudiées à la fin des années 60 par Yves Hellegouarch, en liaison avec le théorème de Fermat, voir [Hellegouarch, 1969 & 1971]. Elles sont revenues sur le devant de la scène en 1985 lorsque Gerard Frey a suggéré que ces courbes elliptiques, si elles existaient, contrediraient la fameuse conjecture STW. La raison qu'il en donnait n'était pas suffisante mais son idée a stimulé de nouvelles recherches (voir en particulier [Serre, 1987]) et a abouti à la démonstration de Ribet. Or la conjecture STW qui a des connexions avec plusieurs autres secteurs des mathématiques, en particulier avec le programme de Langlands évoqué plus haut, apparaissait aux spécialistes beaucoup plus convaincante que le théorème de Fermat lui-même, sur lequel les avis étaient plus partagés, voir [Ribenboim, 1979], chapitre 2. Comme l'a dit Andrew Wiles lui-même dans un interview accordé au journal de son université : "I knew when I heard about Frey and Ribet's result that the lanscape had changed. (...). What Frey and Ribet did was make Fermat's

[24] Pour Taniyama, Shimura et Weil, replacés dans l'ordre alphabétique et non plus chronologique. Elle est connue sous différents noms – en général un sous-ensemble de ces trois – et j'aurai l'occasion de revenir sur les contributions de ces mathématiciens un peu plus tard.

[25] Toutes les courbes de genre 1 ayant au moins un point rationnel, et définies sur $\mathbf{Q}$, admettent une équation de ce type. La quantité Δ s'appelle le discriminant de l'équation. On peut dans notre situation choisir convenablement les coordonnées pour que ce discriminant soit minimal parmi ceux de toutes les équations définissant la même courbe – on parle alors de discriminant de la courbe – et c'est celui-ci que nous noterons Δ par la suite.

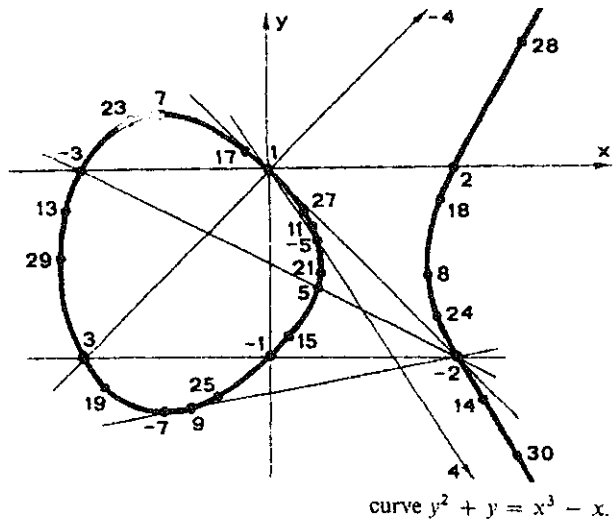
last theorem a consequence of a problem that mathematics could not ignore. Too much depended on it. A whole architecture of future problems depended on it" [26] .

Qu'est-ce donc que cette conjecture "dont toute une architecture dépend" ? En toute première approximation, c'est un analogue pour les courbes de genre 1, définies sur $\mathbf{Q}$ par une équation cubique comme ci-dessus, de la paramétrisation des points du cercle par les fonctions circulaires, cos et sin. Nous allons commencer par donner quelques compléments importants sur les courbes elliptiques, puis nous introduirons les fonctions spéciales (ici les fonctions *modulaires*) qui doivent selon la conjecture intervenir dans leur paramétrisation.

2. Courbes elliptiques (suite).

Par la suite, un point de la courbe E sera pour nous simplement n'importe quel couple (x, y) de nombres *complexes* vérifiant l'équation de E. Si les coordonnées sont réelles, rationnelles, nous parlerons de points réels, rationnels, etc. Par exemple, les points réels de la courbe d'équation $y^2 + y = x^3 - x$ sont dessinés sur le schéma ci-joint et les points marqués **1**, **2**, **3**, etc. sont rationnels. Nous introduirons un point supplémentaire, le point ∞, rencontre des verticales; c'est en fait la solution supplémentaire (rationnelle) qui apparaît lorsqu'on considère non la courbe affine E, mais la courbe projective (dont l'équation générale est

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3).$$



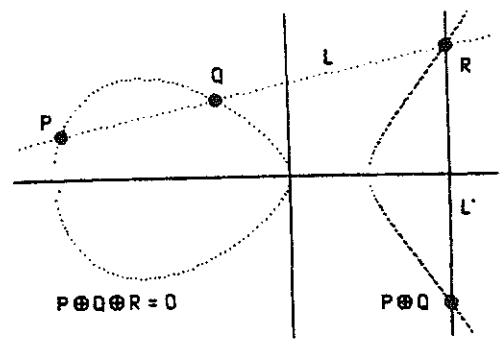
curve $y^2 + y = x^3 - x$.

D'après R. Hartshorne, *Algebraic Geometry*, Springer, 1977

Nous aurons besoin ultérieurement de la propriété cruciale des courbes elliptiques : le fait que l'ensemble de leurs points (resp. de leurs points réels, ou rationnels) forme un groupe commutatif, autrement dit qu'on peut additionner (et soustraire) deux points entre eux. Pour simplifier, nous ne définirons cette addition que lorsque $a_1 = a_3 = 0$ dans l'équation de E (dans ce cas, la courbe est symétrique par rapport à l'axe des abscisses). Partons de deux points quelconques P et Q sur

[26] " J'ai su quand j'ai entendu parler des résultats de Frey et de Ribet que le paysage avait changé. Ce qu'ils avaient fait était de faire du dernier théorème de Fermat la conséquence d'un problème que les mathématiques ne pouvaient pas ignorer. Trop de choses en dépendaient. Toute une architecture de problèmes futurs en dépendait."

la courbe (distincts), la droite qui les joint recoupe en général la cubique en un troisième point R : pour des raisons techniques, ce n'est pas ce point qui est pris comme somme de P et Q mais son symétrique par rapport à l'axe des abscisses. Si $P = Q$, on utilise la tangente à la courbe : le point $P + P = 2P$ est donc le symétrique par rapport à l'axe des abscisses du (troisième) point d'intersection de la tangente en P avec la courbe. L'opposé d'un point P, noté $-P$, est simplement le symétrique de P par rapport à l'axe des abscisses. Le point ∞ est précisément l'élément neutre 0 [27] , par exemple parce que la droite joignant P et son opposé est verticale, donc passe par ∞.



D'après J. Silverman, *The Arithmetic of Elliptic curves*, Springer, 1986

On obtient sans peine les coordonnées de $P + Q$ (ou de $nP = \overbrace{P + P + \cdots + P}^{n \text{ fois}}$) à partir de celles de P et Q et de l'équation de la courbe E, voir par exemple [Silverman] pour des détails. On remarquera que contrairement au groupe des entiers relatifs par exemple, il y a ici des points de torsion, c'est-à-dire des points dont un multiple entier est nul : par exemple, les points d'ordre 2 (vérifiant $P + P = 0$, P non nul) sont exactement les trois points (à coordonnées réelles ou non selon l'équation de la courbe) d'intersection avec l'axe Ox. Les points d'ordre 3 sont les points d'inflexion (le vérifier sur la définition géométrique!). L'étude de certaines permutations de ces points sera cruciale pour la suite.

3. Les fonctions et formes modulaires

Leur définition dépend de deux entiers N (appelé le niveau) et k (le poids; nous n'aurons en fait besoin ici que de k=0 et 2).

Définition : Une fonction modulaire h de niveau N et de poids k est une fonction méromorphe définie sur le demi-plan complexe $H = \{z \in \mathbf{C} \mid \mathrm{Im}(z) > 0\}$, telle que

$$h\left(\frac{az+b}{cz+d}\right) = (cz+d)^k h(z),$$

[27] Pour une preuve, voir par exemple [Reid] ou [Silverman]. Le plus efficace est de s'entraîner sur les dessins (qui ne représentent que les points réels des deux courbes). Lorsque la courbe n'est pas symétrique par rapport à l'axe des abscisses, il faut corriger le processus en conséquence. Dans le cas par exemple de la courbe $y^2 + y = x^3 - x$ dessiné sur le schéma ci-dessus, les points marqués 2, 3 etc, sont obtenus par itération à partir du point marqué 1 : $2 = 1 + 1$, $3 = 1 + 1 + 1$, le $+$ est celui de la loi de groupe que nous venons de définir. On vérifie sur le schéma que par exemple les points -3,2 et 1 sont alignés.

pour toute matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ appartenant au groupe [28]

$$\Gamma_0(N) = \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}) \quad \text{avec} \quad c \equiv 0 \bmod N \}.$$

Rappelons que $SL_2(\mathbf{Z})$ est le groupe des matrices à coefficients entiers et de déterminant 1. Remarquons que la matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_0(N)$ et donc que $h(z+1) = h(z)$: la propriété de la définition généralise bien une notion de période. Etant donc périodique de période 1, les fonctions modulaires ont un développement de Fourier

$$h(z) = \sum_{-\infty}^{\infty} c_n \exp(2\pi i n z),$$

et nous imposerons que ce développement n'ait au plus qu'un nombre fini de coefficients c_n pour $n < 0$, non nuls [29] .

Exemple : La série formelle obtenue en développant le produit

$$q \prod_{n=1}^{\infty} (1-q^n)^2 (1-q^{11n})^2,$$

avec $q = \exp(2\pi i z)$, est une fonction modulaire de poids 2 et de niveau 11. Tous ces coefficients pour $n < 0$ et $n = 0$ sont nuls.

4. La conjecture STW : première formulation

Conjecture : Soit E une courbe elliptique définie sur $\mathbf{Q}$, il existe une paramétrisation par des fonctions modulaires h et g de poids 0 et de niveau N convenable,

$$x = g(z) \quad y = h(z),$$

pour tout point $P = (x, y)$ de E (sauf éventuellement un nombre fini).

Remarques

• Le niveau N peut être défini à partir de la courbe E (c'est ce qu'on appelle son conducteur), voir [Silverman]. Je me contenterai ici de dire que ce conducteur a exactement les mêmes diviseurs premiers que le discriminant minimal de E, Δ. Pour les courbes d'Hellegouarch liées au théorème de Fermat, qui nous intéressent tout particulièrement, le conducteur N est égal [30] à $\prod_{p|abc} p$. Une courbe est dite semi-stable si son conducteur n'a pas de facteurs carrés. Ce phénomène est lié au fait que, lorsqu'on réduit l'équation de E modulo un nombre premier divisant le conducteur

[28] La définition s'étend aux transformations appartenant à un groupe de congruence Γ quelconque : un tel groupe est un sous-groupe d'indice fini de $SL_2(\mathbf{Z})$ contenant le groupe

$$\Gamma(N) = \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}), \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod N \}.$$

Le groupe $\Gamma_0(N)$ du texte est un exemple particulier important de groupe de congruence et nous ne parlerons que de celui-ci ; Wiles en utilise différents autres dans son travail.

[29] Cette condition est en fait une condition de méromorphie à l'infini. D'autres conditions de régularité analogues à celle-ci pour d'autres points doivent être requises également, voir [Lang] ou [Shimura].

[30] Ce conducteur est donc très petit par rapport au discriminant de la courbe. Ceci semble déjà bizarre en soi et des recherches ont été menées pour tirer de là une contradiction, voir [Œsterlé], sans recourir à la conejcture STW.

(donc le discriminant), la courbe obtenue a une singularité ; le cas semi-stable est celui où toutes ces singularités sont des points doubles ordinaires (l'alternative serait la présence de points de rebroussements). La stratégie proposée par Wiles ne vaut en tout cas que pour les courbes semi-stables[31] . Le problème du conducteur (c'est-à-dire du niveau de la forme) est très délicat techniquement : une importante partie du travail de Wiles, et de celui de Ribet avant lui, est lié à la gestion adéquate des niveaux.

• Toute courbe définie par une équation cubique comme précédemment, à coefficients *complexes*, admet une paramétrisation par des fonctions périodiques, les fonctions de Weierstrass ; c'est un résultat classique bien connu dès le 19ème siècle. Mais que des fonctions *modulaires* pour un sous-groupe de congruence comme ci-dessus fournissent une (autre) paramétrisation est étroitement lié au fait que les coefficients de l'équation cubique soient *rationnels* : la conjecture STW est une conjecture arithmétique[32] .

Exemple. Une paramétrisation modulaire (par des fonctions de poids 0 et de niveau N=37) pour la courbe $y^2 + y = x^3 - x$ est donnée par[33] :

$$x(z) = q^{-2} + 2q^{-1} + 5 + 9q + 18q^2 + 29q^3 + \cdots$$

$$y(z) = q^{-3} + 3q^{-2} + 9q^{-1} + 21 + 46q + 92q^2 + \cdots,$$

avec comme toujours $q = \exp(2\pi iz)$.

5. D'autres formulations de la conjecture STW

Cette formulation relativement simple n'est pas la seule (ni la plus courante si l'on regarde la littérature sur le sujet). La conjecture établit en fait un dictionnaire entre des objets "elliptiques" (définis à partir de la courbe E) et des objets "modulaires" (définis à partir de groupes de congruences et de fonctions modulaires). Je vais donner ici plusieurs de ces correspondances : le fait que la conjecture STW puisse être prouvée à partir de l'une d'entre elles, "au choix", est souvent loin d'être immédiat et repose sur les travaux de nombreuses personnes.

• Interprétation géométrique

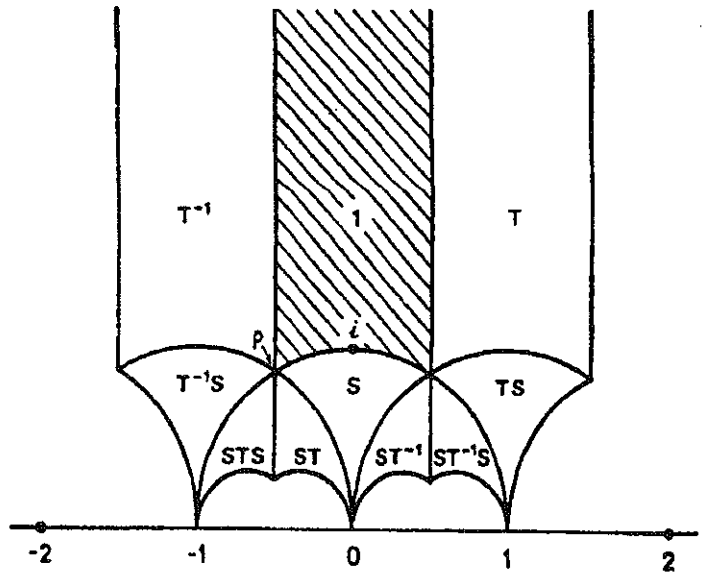
Pour un entier N fixé, le groupe $\Gamma_0(N)$ définit un pavage du demi-plan supérieur : on peut en effet découper ce plan en une infinité de régions qui se déduisent les unes des autres par une transformation homographique du type $z \to \frac{az+b}{cz+d}$, avec $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$. Une fonction modulaire de poids 0 étant invariante par une telle transformation, elle est définie partout dès qu'on la connaît sur une seule de ces régions. Les bords d'une telle région (sauf un nombre fini de points) peuvent aussi être recollés en identifiant ceux qui ne diffèrent que par une transformation de $\Gamma_0(N)$. Nous dessinons ci-après le cas de $N = 1$. La transformation S est $z \to -1/z$, la transformation T est la translation par 1. On a indiqué sur chaque région quelle transformation la déduirait de la région marquée **1**.

---

[31] Rappelons pour mémoire que sa démonstration même dans ce cas n'est pas complète (mars 1994).

[32] Pour de plus amples informations là-dessus, voir [Mazur].

[33] Le détail des calculs figure dans [Mazur & Swinnerton-Dyer] ou [Zagier].



Domaine fondamental des fonctions modulaires de niveau 1

On obtient ainsi une surface réelle (qu'on compactifie en ajoutant les points manquants) qu'on appelle usuellement $X_0(N)$. C'est une surface de Riemann. Les fonctions modulaires de poids 0 sont simplement les fonctions sur cette surface. L'existence d'une paramétrisation modulaire pour une courbe elliptique E définie sur $\mathbf{Q}$ revient à l'existence d'une application holomorphe non constante entre surfaces de Riemann (ou encore d'une application rationnelle entre courbes algébriques)[34] :

$$X_0(N) \to E.$$

• Formes modulaires de poids 2

Une forme très importante de la conjecture est qu'il est possible d'associer à E une fonction modulaire f de poids 2 et de niveau N avec les propriétés suivantes :

i) c_1=1, $c_n = 0$ pour $n \leq 0$ (une telle fonction est dite : forme modulaire parabolique normalisée).

ii) Pour tout nombre premier l ne divisant pas N,

$$c_l = 1 + l - \text{card}(E \bmod l).$$

Autrement dit, le l-ième coefficient de Fourier de la forme modulaire s'obtient en déduisant de $l+1$ le nombre de solutions modulo l de l'équation de E réduite modulo l (attention, il faut aussi compter la solution ∞ dans la réduction). La formule (ii) est fondamentale et nous la retrouverons souvent par la suite.

[34] Peut-être n'est-il pas inutile de rappeller au passage qu'une surface de Riemann est bien "une surface" à 2 dimensions *réelles*, mais que c'est aussi une courbe *complexe*. Le cas le plus évident est bien sûr $\mathbf{C}$ lui-même, droite complexe repérée par le paramètre z et plan réel repéré par les deux coordonnées Re(z), Im(z) (Attention, avec mes notations, les coordonnées x et y utilisées pour écrire l'équation de la courbe E n'ont bien sûr rien à voir avec les parties réelles et imaginaires de la coordonnée complexe z !). En l'occurrence, on peut prouver que la courbe $X_0(N)$ a une équation définie sur $\mathbf{Q}$; $X_0(37)$, par exemple, est, comme surface, un double pneu, et comme courbe, l'ensemble des solutions (complexes) de l'équation $V^2 = 37 - U^6 - 9U^4 - 11U^2$, dans un plan de coordonnées U et V.

Remarques : Les formes modulaires de poids 2 donnent lieu à des formes différentielles sur la surface $X_0(N)$; celle associée à E est l'image inverse par l'application $X_0(N) \to E$ de la forme différentielle $\frac{dx}{2y+a_1}$ sur E. L'espace des formes modulaires paraboliques de poids 2 et de niveau N est un espace vectoriel de dimension finie : 0 pour N=2, 1 pour N=11, 2 pour N=37, par exemple. Cette dimension est comme il se doit le genre de la courbe $X_0(N)$.

La forme f associée à E doit vérifier d'autres propriétés importantes : c'est un vecteur propre sous l'action d'opérateurs, dont les T_n ($0 \neq n \in \mathbf{N}$), dits "opérateurs de Hecke", qui interviennent de façon cruciale dans les définitions (si on veut les rendre précises...) et dans les preuves, voir par exemple [Lang]. Les coefficients c_n sont simplement les valeurs propres correspondantes à chaque T_n. Des formules de récurrence simples permettent de déterminer tous les coefficients c_n dès que les c_l, l premier, sont connus.

Exemple. Nous avons déjà donné une paramétrisation pour la courbe $y^2 - y = x^3 - x$, on en déduit une forme modulaire, parabolique normalisée, par

$$f(z)dz = \frac{dx(z)}{2y(z) - 1},$$

soit le développement :

$$f(z) = q - 2q^2 - 3q^3 + 2q^4 - 2q^5 + 6q^6 + \cdots,$$

avec $q = \exp(2\pi iz)$ comme toujours. On peut vérifier la propriété (ii) sur les premiers coefficients. Par exemple, les solutions modulo 3 de l'équation de la courbe sont : (0,0), (0,1), (0,2), (1, 0), (1,1), (1, 2), soit 7 solutions avec celle à l'infini, donc on doit avoir $c_3 = 3 + 1 - 7 = -3$ comme annoncé.

• Fonctions L

A une forme modulaire f parabolique de poids 2 et de niveau N comme précédemment, on peut associer sa transformée de Mellin

$$\int_0^\infty f(it)t^s \frac{dt}{t},$$

qui est encore égale à

$$(2\pi)^{-s}\Gamma(s)L(f,s),$$

avec $\Gamma(s) = \int_0^\infty e^{-t}t^s \frac{dt}{t}$ et $L(f,s) = \sum c_n n^{-s}$. Cette série L converge pour $Re(s) > 3/2$ et admet un prolongement analytique à tout le plan complexe ; elle vérifie de plus une équation fonctionnelle reliant sa valeur en s à sa valeur en $2-s$.[35] La conjecture STW peut encore être exprimée en disant que la fonction L de la courbe elliptique, définie à partir du nombre de points des réductions de E modulo les nombres premiers, coïncide avec une fonction $L(f,s)$. Voici comment est construite la fonction L de la courbe elliptique. Pour un nombre premier l ne divisant pas le niveau N, on pose

$$L_l(E,s) = (1 - (1 + l - \mathrm{card}(E \bmod l))l^{-s} + l^{1-2s})^{-1}.$$

La fonction $L(E,s)$ de la courbe elliptique est le produit des $L_l(E,s)$ pour tous les nombres premiers l (y compris des facteurs adéquats que je ne définis pas ici pour les diviseurs de N). Elle

[35] Ces propriétés rappellent bien sûr celle de la célèbre fonction ζ de Riemann, dont l'équation fonctionnelle lie la valeur en s à celle en $1 - s$.

détermine la courbe E à isogénie près, c'est-à-dire qu'une courbe ayant même fonction L que E est un quotient de E par un groupe fini[36] .

André Weil a prouvé qu'il suffit de montrer que la fonction L de la courbe (et une famille d'autres déduites de celle-ci) admet un prolongement analytique et vérifie une équation fonctionnelle du type voulu pour en déduire la conjecture STW (voir [Weil, OS], tome 3, p.165 pour l'énoncé exact). Cet énoncé a beaucoup contribué à rendre vraisemblable la conjecture STW, car on attendait en principe de tels prolongements et équations fonctionnelles pour les fonctions L associées à des courbes ou à des variétés algébriques – remarquons toutefois qu'on ne sait pas les démontrer a priori pour les courbes elliptiques définies sur $\mathbf{Q}$ sauf hypothèses supplémentaires.

Remarque : Il semble que ce soit le mathématicien Taniyama qui ait le premier proposé, lors d'une session de problèmes à la fin d'une conférence à Tokyo en 1955, une version faible de STW : selon la version anglaise apparemment non publiée, mais "diffusée assez largement", comme on dit, Taniyama demandait s'il était possible de trouver une "forme automorphe" (a priori quelque chose d'un peu plus général que les fonctions modulaires) dont la transformée de Mellin donnerait la fonction L de la courbe elliptique. Dans une série d'articles, Goro Shimura construisit en particulier pour chaque forme modulaire de poids 2 (de niveau N, parabolique, etc. comme ci-dessus) dont le développement en série de Fourier est à coefficients rationnels, une courbe elliptique associée – si les coefficients appartiennent à une extension de degré fini, on obtient des variétés de dimension supérieure – et montra que ces courbes étaient modulaires au sens de l'interprétation géométrique ci-dessus. Ceci établit une moitié du dictionnaire entre "elliptique" et "modulaire". Weil comme nous l'avons dit a prouvé que la propriété de modularité pour une courbe elliptique découlerait de propriétés simples de fonctions L associées. L'intérêt du dictionnaire est que les fonctions modulaires sont plus concrètes, plus accessibles que les courbes elliptiques : par exemple, on sait prouver les propriétés de prolongement de $L(f,s)$. De même, puisque la dimension de l'espace des formes modulaires de poids 2 et de niveau N est finie, la conjecture STW entraîne par exemple qu'il n'y a qu'un nombre fini de courbes elliptiques de conducteur donné.

6. Représentations galoisiennes

Nous allons donner une formulation supplémentaire qui est fondamentale dans les travaux de Ribet et Wiles. Elle repose sur l'étude des représentations d'un même groupe, le groupe de Galois absolu, dans différents modules, attachés soit aux courbes elliptiques, soit aux formes modulaires.

• Le groupe de Galois absolu

Le groupe de Galois d'une équation polynomiale à coefficients rationnels est simplement le groupe des permutations des racines de cette équation qui conservent toutes les relations entre elles s'exprimant par des polynômes à coefficients rationnels. Autrement dit, si les racines $x_1, \cdots, x_n$ d'une équation vérifient $Q(x_i) = 0$, où Q est un polynôme à coefficients rationnels, les images $s(x_i)$, pour une permutation s du groupe de Galois, donnent encore vérifier $Q(s(x_i)) = 0$. En particulier, le groupe de Galois fixe toutes les racines rationnelles (si $x_1 = a$ avec a rationnel, $s(x_1)$ doit aussi être égal à a).

Par exemple, soit l'équation $x^4 - x^3 - 2x^2 + 3x - 1 = 0$. Elle a quatre racines, $x_1 = 1, \quad x_2 =$

[36] On ne sait pas prouver en général que ces fonctions $L(E,s)$ ont un prolongement analytique à tout le plan complexe. Ceci se déduirait de la conjecture STW puisqu'on le sait déjà pour $L(f,s)$. La célébrissime conjecture de Birch et Swinnerton-Dyer – encore une !– affirme que le comportement de la fonction $L(E,s)$ au voisinage de 1 est lié à la structure des points *rationnels* de la courbe.

-1, $x_3 = \frac{-1+\sqrt{5}}{2}$, $x_4 = \frac{-1-\sqrt{5}}{2}$. Il y a a priori $4! = 24$ permutations entre les 4 racines, mais celles autorisées doivent conserver par exemple les relations $x_1 - 1 = 0$, $x_2 + 1 = 0$, $x_3 + x_4 + 1 = 0$, $x_3 x_4 - 1 = 0$. Elles fixent donc x_1 et x_2, et fixent les deux autres racines ou les échangent entre elles. Le groupe de Galois contient donc deux éléments, l'identité, et la permutation qui échange x_3 et x_4 : cette dernière correspond à la transformation $\sqrt{5} \to -\sqrt{5}$.

Soit par contre l'équation $x^4 + x^3 + x^2 + x + 1 = 0$, dont les racines sont les racines 5-ièmes de l'unité, différentes de 1. Chaque racine est une puissance de l'une d'elles, par exemple de x_1, et cette relation doit être conservée par les permutations autorisées : une telle permutation est déterminée dès qu'on connait l'image de x_1 : il y a quatre possibilités (x_1, x_2, x_3, x_4) et le groupe de Galois a quatre éléments, c'est en l'occurrence un groupe cyclique d'ordre 4.

Soit enfin l'équation $3x^4 - 6x^2 + x - 1 = 0$, il n'y a pas de relations polynomiales entre les racines sauf l'équation elle-même, le groupe de Galois de l'équation est ici le groupe S_4 de toutes les permutations entre les 4 racines.

Le groupe de Galois absolu, $G_{\mathbf{Q}}$, réunit toutes ces informations pour toutes les équations polynomiales à coefficients rationnels : un élément du groupe de Galois absolu détermine ainsi sur chaque équation une permutation particulière des racines (qui peut bien entendu être l'identité). Pour plus de détails, voir par exemple [Stewart]. Ce groupe de Galois intéresse beaucoup les arithméticiens, car les méthodes de géométrie algébrique marchent surtout (mais pas exclusivement !) sur des corps algébriquement clos : le groupe de Galois contrôle le passage de la clôture algébrique de $\mathbf{Q}$ (contenant les racines des équations polynomiales à coefficients rationnels) à $\mathbf{Q}$ lui-même.

Nous allons maintenant interpréter ce groupe de Galois comme un groupe de transformations de modules (ou espaces vectoriels[37]) appropriés, associés, soit à une courbe elliptique, soit à une forme modulaire. La conjecture STW affirmera que ces interprétations sont essentiellement les mêmes.

• Représentations galoisiennes elliptiques

Soit n un entier > 2 et $E[n] = \{P \mid \overbrace{P + \cdots + P}^{n \text{ fois}} = 0\}$ – autrement dit, l'ensemble des points, à coordonnées complexes, de la courbe E dont le n-ième multiple est 0. On inclut le point à l'infini (élément "zéro" pour la loi de groupe) et les $E[n]$ forment des groupes commutatifs de n-torsion, autrement dit des $\mathbf{Z}/n\mathbf{Z}$-modules. Il n'est pas très difficile, à partir de l'expression des coordonnées de la somme de plusieurs points, de constater que les abscisses x des points de $E[n]$ vérifient une équation de degré $(n^2 - 1)/2$ à coefficients rationnels; les ordonnées sont données par l'équation de E (deux valeurs de y pour chaque x). Il y a donc en tout n^2 points dans $E[n]$ (en comptant le point ∞); en fait, $E[n]$ forme un $\mathbf{Z}/n\mathbf{Z}$-module libre de rang 2.

Le groupe de Galois absolu conserve les relations algébriques, en particulier l'équation aux abscisses des points de $E[n]$ et les équations donnant les ordonnées pour une abscisse fixée. Donc il permute les points de $E[n]$ entre eux. Fixons une base (P_1, P_2) du $\mathbf{Z}/n\mathbf{Z}$-module $E[n]$: pour chaque élément σ du groupe de Galois, $\sigma(P_1)$ et $\sigma(P_2)$ s'expriment dans cette base, comme une combinaison linéaire à coefficients dans $\mathbf{Z}/n\mathbf{Z}$: $\sigma(P_i) = r_i P_1 + s_i P_2$. On obtient ainsi une représentation[38]

[37] Il y a de bonnes raisons de préférer parfois les modules sur un anneau convenable et parfois les espaces vectoriels sur le corps des fractions; on ignorera cette question par la suite.

[38] C'est un homomorphisme continu de groupes topologiques, je ne décris pas ici les topologies nécessaires.

$$\rho_n : G_{\mathbf{Q}} \longrightarrow GL_2(\mathbf{Z}/n\mathbf{Z}),$$

en associant à chaque élément σ de $G_{\mathbf{Q}}$ la matrice $\begin{pmatrix} r_1 & r_2 \\ s_1 & s_2 \end{pmatrix}$.

Exemple : Pour la courbe $y^2 - y = x^3 - x$ et $n = 3$, l'équation donnant les abscisses des points de $E[3]$ est : $3x^4 - 6x^2 + x - 1 = 0$, le groupe de Galois de l'équation est ici le groupe S_4. L'action du groupe de Galois absolu est déterminée par celle sur les abscisses et celle sur les ordonnées des points. On pourrait décrire explicitement la représentation ρ_3 ci-dessus en choisissant une base de $E[3]$ sur $\mathbf{Z}/3\mathbf{Z}$.

On a une application naturelle de $E[n]$ dans $E[m]$, dès que m divise n (si $P \in E[n]$, alors $(n/m)P \in E[m]$) et les représentations associées sont compatibles entre elles. En particulier, il est possible de définir une unique représentation galoisienne réunissant toutes les informations contenues dans les représentations galoisiennes sur $E[3]$, $E[9]$, $E[27]$, etc., (ou les analogues pour des puissances d'un autre nombre premier à la place de 3). Cette représentation est dite 3-adique, nous la noterons ρ_{3^∞},

$$\rho_{3^\infty} : G_{\mathbf{Q}} \longrightarrow GL_2(\mathbf{Z}_3),$$

où $\mathbf{Z}_3$ est l'anneau des entiers 3-adiques [39] .

Un résultat crucial est que la représentation 3-adique (la même chose serait vraie en remplaçant 3 par presque tout autre nombre premier) suffit à déterminer la fonction L de la courbe (donc la courbe E elle-même à isogénie près).

• Représentations modulaires

Il s'agit de déterminer le même type d'objets à partir d'une forme modulaire (de poids 2, de niveau N, parabolique, etc). Je ne dirai pas exactement comment réaliser concrètement, sur un module précis, les représentations voulues (ceci fait par exemple intervenir les opérateurs de Hecke évoqués plus haut et est assez technique, voir [Serre] ou [Ihara, Ribet & Serre]), et me contenterai d'en décrire une propriété caractéristique.

Comme on peut le deviner, les coefficients c_l, pour les nombres premiers l, jouent un rôle fondamental. Pour tout nombre premier l, on peut définir un élément privilégié dans le groupe de Galois absolu, qu'on note σ_l [40] . On associe alors à une forme f une représentation

$$\rho_{3^\infty} : \quad G_{\mathbf{Q}} \longrightarrow GL_2(\mathbf{Z}_3),$$

telle que la trace de la matrice $\rho_{3^\infty}(\sigma_l)$ soit exactement le coefficient c_l (pour l ne divisant pas $3N$). On pourrait définir également des représentations ρ_3, ρ_9, ρ_{27}, etc., les traces correspondantes étant respectivement c_l modulo 3, c_l modulo 9, c_l modulo 27, etc. Une construction analogue vaut si on remplace 3 par presque tout autre nombre premier, mais comme on le verra, le cas de 3 est celui qui est crucial pour Wiles.

---

[39] Un nombre 3-adique est analogue à une série entière infinie en les puissances de 3, à coefficients dans $\mathbf{Z}/3\mathbf{Z}$: $a_0 + a_1 3 + a_2 3^2 + \cdots$. On obtient les renseignements modulo 3^r en tronquant le développement à la puissance voulue. En particulier, si $\rho_{3^\infty}(\sigma) = \begin{pmatrix} r_1 & r_2 \\ s_1 & s_2 \end{pmatrix}$, alors $\rho_{3^r}(\sigma) = \begin{pmatrix} r_1 \bmod 3^r & r_2 \bmod 3^r \\ s_1 \bmod 3^r & s_2 \bmod 3^r \end{pmatrix}$.

[40] En fait une classe d'éléments, les automorphismes de Frobenius en l, dont la définition exigerait d'autres informations sur la structure du groupe de Galois, voir par exemple [Cassels & Frohlich].

• La conjecture de STW s'exprime alors en disant que pour toute courbe elliptique E définie sur $\mathbf{Q}$, la représentation $\rho_{3\infty}$ associée à E est isomorphe à une représentation provenant d'une forme modulaire comme ci-dessus. Notons que c'est $\rho_{3\infty}$ toute entière qui doit être modulaire, la modularité de l'une des représentations ρ_{3^r} ne suffirait pas.

Dans la suite, nous noterons $\rho_{n,E}$, $\rho_{3\infty,E}$, etc., les représentations construites à partir d'une courbe elliptique, et $\rho_{n,f}$, $\rho_{3\infty,f}$, etc. celles construites à partir des coefficients d'une forme modulaire de poids 2, parabolique, etc. Cette formulation est en particulier compatible avec la formule (ii) du paragraphe précédent.

Retour sur le programme de Langlands. On associe en fait une fonction L à toute représentation galoisienne convenable. Le programme de Langlands prédit que ces fonctions L coïncident avec les fonctions L de formes automorphes, qui généralisent les formes modulaires.

Retour sur la notion de représentation galoisienne. Il y a deux points de vue possibles sur ces représentations. Le premier s'intéresse plutôt au groupe de Galois lui-même : les différents modules ou espaces vectoriels d'arrivée sont autant de manières de l'étudier. Le second est orienté vers la connaissance des modules ou espaces vectoriels eux-mêmes, et on pense à la représentation comme une structure supplémentaire sur eux. On a tendance dans ce cas, qui est le nôtre, à laisser tomber le graphisme

$$\rho_{3,E} : \quad G_{\mathbf{Q}} \longrightarrow GL(E[3]) = GL_2(\mathbf{Z}/3\mathbf{Z}),$$

pour parler simplement de $E[3]$ comme $G_{\mathbf{Q}}$-module de rang 2 (sous-entendu : $E[3]$ muni de la structure supplémentaire que lui procure la représentation ρ_3 de $G_{\mathbf{Q}}$. De même pour $\rho_{3\infty}$). Il ne s'agit que d'une façon de parler ; elle a pour moi l'avantage que les travaux de Wiles vont dans la suite faire apparaître d'autres $G_{\mathbf{Q}}$-modules de rang 2, modules sur des anneaux plus compliqués que $\mathbf{Z}/3\mathbf{Z}$ ou $\mathbf{Z}_3$. Ce point de vue m'évitera d'avoir à les préciser.

7. Les résultats de Ribet

Il est temps maintenant de revenir brièvement au lien entre la conjecture STW et le théorème de Fermat. On prend bien sûr comme courbe E une courbe de Hellegouarch d'équation $y^2 = x(x-a^p)(x-b^p)$, où $a^p + b^p = c^p$ vérifient la relation de Fermat. Si elle était modulaire, comme le prévoit la conjecture STW, les représentations $\rho_{n,E}$ pour tout entier n, seraient modulaires au sens précédent. Ce serait en particulier le cas pour[41] $\rho_{p,E}$. Autrement dit, pour tout nombre premier l qui ne divise pas pN,

$$\mathrm{Tr}\rho_{p,E}(\sigma_l) \equiv c_l \,\mathrm{modulo}\, p,$$

où les c_l sont les coefficients de Fourier d'une forme modulaire de niveau N.

Ce qu'a montré Ribet, c'est que pour cette représentation, il serait possible de baisser le niveau,[42] plus précisément qu'il existerait une forme modulaire de poids 2 et de niveau 2, $\sum c'_n q^n$,

[41] Attention, ce n'est pas une faute typographique ! Le p est bien l'exposant de l'équation de Fermat qu'on considère. Le problème des notations dans ce sujet n'est pas tout à fait évident : il était d'usage d'avoir deux nombres premiers en jeu, l'un fixe (indice de la représentation), l'autre variable (lorsqu'on décrit les coefficients de Fourier). La tradition notait le dernier p, le premier fixe l. Avec l'arrivée des courbes de Hellegouarch, p s'est trouvé capturé comme exposant de l'équation de Fermat associée. Ici heureusement, l'indice de la représentation utile est soit p lui-même chez Ribet, soit 3 ou 5 chez Wiles. Mais ce n'est pas toujours le cas. Cherche lettre désespérément. Naturellement, $q = exp(2\pi iz)\ldots$

[42] Le jeu, très compliqué, consiste à éliminer un par un les facteurs impairs du conducteur. Voir [Œsterlé] et [Ribet,

telle que

$$\mathrm{Tr}\rho_{p,E}(\sigma_l) \equiv c'_l \equiv c_l \,\mathrm{modulo}\, p$$

pour presque tout nombre premier l. Or il n'existe pas de forme modulaire parabolique de poids 2 et de niveau 2, d'où une contradiction. Ceci prouve que si la courbe d'Hellegouarch existait (c'est-à-dire si le théorème de Fermat était faux pour un p), elle ne pourrait être modulaire : la conjecture STW implique donc le théorème de Fermat.

Elle impliquerait d'ailleurs tout aussi bien d'autres résultats du même type, voir [Serre]. Par exemple, le théorème de Ribet montre qu'on dédui(rai)t de la conjecture STW que les équations

$$a^p + b^p + L^\alpha c^p = 0,$$

pour $\alpha \geq 0$ et $L \in \{3,5,7,11,13,17,19,23,29,53,59\}$, $L \neq p$, n'ont pas de solutions entières non nulles.

8. Vers une preuve de la conjecture STW pour les courbes elliptiques semi-stables : les travaux d'Andrew Wiles

Tout ceci était donc en place avant le récent travail de Wiles. Nous sommes maintenant en mesure de donner une idée des étapes principales qu'il a suggéré de parcourir pour prouver la conjecture STW pour toutes les courbes elliptiques définies sur $\mathbf{Q}$ semi-stables. Partons donc d'une telle courbe elliptique. Il s'agit de montrer qu'elle est modulaire, par exemple en prouvant que la représentation elliptique $\rho_{3^\infty,E}$ est modulaire.

• La première étape consiste à montrer que, dans beaucoup de cas, $\rho_{3,E}$ est modulaire, autrement dit qu'il existe une forme modulaire f de niveau N et de poids 2, $\sum c_n q^n$, telle que

$$\mathrm{Tr}(\rho_{3,E}(\sigma_l)) \equiv \quad c_l \quad \text{modulo} \quad 3,$$

pour tout l ne divisant pas $3N$.

L'existence d'une telle forme provient de ce que, dans les cas où le groupe de Galois de l'équation aux abscisses de $E[3]$ est tout le groupe S_4 des permutations des 4 racines (comme c'est le cas par exemple pour la courbe $y^3+y = x^3-x$), la représentation $\rho_{3,E}$ est une représentation pour laquelle on sait prouver une portion du programme de Langlands (travaux de Langlands et Tunnell [43]) ; des résultats de Deligne et Serre permettent d'en déduire la forme modulaire f de poids 2 cherchée. Il faut remarquer que le programme de Langlands est connu dans très peu de cas : le fait de choisir 3 (pour lequel l'équation aux abscisses est de degré 4) est crucial pour que la preuve puisse commencer.

• On dispose maintenant de deux représentations ρ_{3^∞}, l'une est celle associée à E, l'autre à la forme f obtenue dans l'étape précédente (elles ne coïncident que modulo 3). Le principe consiste alors à utiliser une théorie décrivant les relèvements des représentations modulo 3 aux représentations 3-adiques (ou selon l'optique alternative dont j'ai parlé, des $\mathbf{Z}/3\mathbf{Z}$-modules galoisiens de rang 2 aux $\mathbf{Z}_3$-modules galoisiens de rang 2) ; cette théorie a été élaborée principalement ces dernières années par Mazur, Hida, Tilouine. Wiles a construit une représentation galoisienne

1990 a & b].

[43] On trouvera les nombreuses références de cette section explicitées dans [Ribet, 1993] ou [Goldstein, 1994].

modulaire "universelle" telle que toutes les représentations modulaires relevant $\rho_{3,f}$ s'en déduisent naturellement (ou dans notre optique : un module galoisien $\mathbf{T}$ sur un anneau adéquate, qui soit "modulaire universel", au sens où tous les modules galoisiens sur $\mathbf{Z}_3$ associés à des formes modulaires congrues à f modulo 3, en particulier à f elle-même, soient des images de $\mathbf{T}$ par des applications convenables). Il a construit de même une représentation "géométrique universelle", $\mathbf{U}$ qui joue le même rôle pour les représentations d'origine géométrique [44] relevant $\rho_{3,E}$, en particulier $\rho_{3^\infty,E}$.

• La dernière étape consiste à comparer les deux types de relèvements et à montrer que $\mathbf{T}$ et $\mathbf{U}$ coïncident, ce qui permettrait de conclure que $\rho_{3^\infty,E}$ est modulaire. Cette partie est la plus délicate. Elle contient un examen algébrique de la structure de ces objets universels, puis une estimation de l'indice de l'un dans l'autre. Cet indice est interprété comme un analogue géométrique d'un nombre de classes et Wiles a tenté de le calculer en s'inspirant d'idées de Thaine, Kolyvagin, Rubin, Flach, qui avaient traité depuis 1986 des cas plus simples. Cet indice est à peu près l'ordre d'un groupe, un certain "groupe de Selmer", comme il se doit dans le cadre de l'analogie déjà évoquée, et, pour l'estimer, il faut construire à la force du poignet des éléments du groupe ayant le bon ordre. Wiles utilise pour cela des résultats délicats de géométrie algébrique. C'est là que situe, semble-t-il, le principal obstacle à la démonstration complète. Selon le message électronique diffusé par Wiles en décembre 1993 : "During the review process a number of problems emerged, most of which have been resolved, but one in particular I have not yet settled. The key reduction of (most cases of) the Taniyama-Shimura conjecture to the calculation of the Selmer group is correct. However the final calculation of a precise upper bound for the Selmer group in the semi-stable case (...) is not yet complete as it stands" [45] .

• Lorsque la représentation $\rho_{3,E}$ n'est pas adéquate pour attraper immédiatement une forme modulaire comme dans la première étape, Wiles a eu recours à la représentation $\rho_{3,E'}$ pour une autre courbe E' auxiliaire et à une comparaison entre $\rho_{5,E}$ et $\rho_{5,E}$ pour conclure.

Remarque : Même si la preuve n'est pas complète pour toutes les courbes elliptiques semi-stables (en particulier pour celles liées au théorème de Fermat), Wiles semble dès maintenant capable de prouver la conjecture STW pour un nombre infini de classes d'isomorphismes (sur $\mathbf{C}$) de courbes elliptiques : avant son travail on ne savait la prouver que lorsque l'équation de la courbe était donnée numériquement ou pour une famille de courbes très particulières, formant un nombre fini de classes d'isomorphismes.

Conclusion

Comme cette très superficielle présentation le laisse soupçonner, des techniques et des concepts venant de plusieurs branches de la géométrie algébrique et de la théorie des nombres contemporaines sont mis à l'œuvre. On a vu au passage comment des idées de la théorie algébrique des nombres classique, par exemple la liaison entre valeurs de fonctions analytiques et structures, occupaient une

[44] Cette terminologie vague recouvre des caractérisations techniques très fines sur les relèvements "géométriques" licites, en particulier pour contrôler la situation pour les diviseurs de 3N. L'examen des propriétés utiles, en liaison avec la géométrie des variétés, est dû entre autres à Bloch-Kato, Fontaine, Mazur, etc. Se trouve caché ici un important travail conceptuel entamé cette décennie.

[45] "Durant l'examen par les rapporteurs plusieurs problèmes sont apparus, dont la plupart ont été résolus, mais un en particulier n'est pas encore réglé. La réduction fondamentale de (la plupart des cas de) la conjecture de Taniyama-Shimura au calcul du groupe de Selmer est correct. Mais le calcul final d'une borne supérieure précise pour le groupe de Selmer dans le cas semi-stable n'est pas complet tel qu'il est". Peut-être faut-il ajouter que, selon les spécialistes, la valeur de l'indice cherché est compatible avec, voire prédit par, les conjectures courantes.

place importante : à bien des égards, le cas des courbes elliptiques est un analogue, en dimension supérieure, des corps cyclotomiques et cette analogie, mêlée à d'autres, a inspiré de nombreuses recherches pendant les dernières décennies, voir [Coates & Taylor] et [Ihara, Ribet & Serre]. Des conjectures multiples prédisent des liens encore mal compris entre différents objets géométriques ou arithmétiques et constituent l'arrière-plan de ce travail : celles de Langlands, de Bloch-Kato, etc. Et il aurait fallu citer beaucoup d'autres noms.

Que cette approche au théorème de Fermat soit couronnée d'un complet succès ou non, elle offre en tout cas quelques perspectives passionnantes. Pour ceux et celles qui, malgré mes efforts, persisteraient à s'intéresser au théorème de Fermat lui-même, je les renvoie à [Frey] où un réseau de conjectures et d'implications le cerne comme à l'hallali. Plus qu'un (tout petit) effort ?

TOLOSÆ,

Apud JOANNEM PECH, Comitiorum Fuxenſium Typographum, juxta Collegium PP. Societatis JESU.

M. DC. LXXIX.

Vignette du 2ème tirage des Varia Opera de Fermat, chez le même éditeur

Bibliographie

L. Bucciarelli et N. Dwork, *Sophie Germain. An essay in the history of the theory of elasticity*,Reidel, Dordrecht, **1980**

J. W. S. Cassels et A. Fröhlich, éd., *Algebraic Number Theory*, Academic Press, London, **1967**

G. Cifoletti, *Mathematics and Rhetoric, Jacques Pelletier, Guillaume Gosselin and the Making of the French Algebraic Tradition*, PhD, Princeton, **1992**

J. Coates et M.J Taylor, éd., *L-functions and arithmetic*, London Mathematical Society LNS 153, Cambridge University Press, Cambridge, **1991**

G. Cornell et J. Silverman, *Arithmetic Geometry*, Springer, Berlin, New York, Heidelberg, **1986**

H. Davenport, *The Higher Arithmetic*, Cambridge University Press, Cambridge, etc., 5-ème édition **1982** (il existe maintenant une 6-ème édition)

J. M. Deshouillers, " Théorème de Fermat : la contribution de Fouvry ", *Séminaire Bourbaki*, 37, **1984-1985**, 648.

L. E. Dickson, *History of the Theory of Numbers*, Volume II : Diophantine Analysis, 1921, reprinted : Chelsea, New York, **1952**

H.M. Edwards, "The Background of Kummer's proof of Fermat's Last Theorem", *Archive for History of Exact Sciences*, 14, 3, **1975**, 219-236

H.M. Edwards, "Postscript to : The Background of Kummer's proof of Fermat's Last Theorem", *Archive for History of Exact Sciences*, 17, **1977**, 381-394

H. M. Edwards, *Fermat's Last Theorem*,Springer, New York Berlin Heidelberg, **1977**

H. Edwards, "Mathematical Ideas, Ideals and Ideology", *Mathematical Intelligencer*, 14, 2, **1992**, 6-18

P. de Fermat, *Œuvres*, 4 volumes (plus un Supplément de C. de Waard), éditées par P. Tannery et C. Henry, 1896-1912

G. Frei, " Heinrich Weber and the Emergence of Classe Field Theory ", *The History of Modern Mathematics*, D. Rowe et J. Mc Cleary éd., Academic Press, Boston, etc., **1989**, 425-450

G. Frey, "Links between solutions of A-B=C and elliptic curves ", *Number Theory, Proceedings of the Journées Arithmétiques - Ulm, 1987*, Lecture Notes in Mathematics, 1380, Springer, **1989**

S. Gelbart, "An elementary introduction to the Langlands program", *Bulletin of the AMS*,10, 2, **1984**, 177-219

C. Goldstein, "Le métier des nombres au 17 ème et au 19 ème siècles", *Eléments d'Histoire des Sciences*, M. Serres, éd., Bordas, Paris, **1989**

C. Goldstein, "Descente infinie et analyse diophantienne : programmes de travail et mises en œuvre chez Fermat, Levi, Mordell et Weil", *Cahier du Séminaire d'Histoire des Mathématiques*, 2, 3, **1993**, 25 sq.

C. Goldstein, "Le théorème de Fermat", *La Recherche*, **mars 1994**

C. Goldstein a, "La théorie des nombres dans les Notes aux Comptes Rendus de l'Académie des Sciences (1870-1914)", **1994**, à paraître

C. Goldstein b, *L'aire d'un triangle rectangle en nombres n'est pas un carré : les lectures d'un résultat mathématique*, **1994**, à paraître.

Y. Hellegouarch, *Courbes elliptiques et équations de Fermat*, Thèse d'Etat, Université de Besançon, **1969**

Y. Hellegouarch, *Notes aux Comptes Rendus de l'Académie des Sciences*, 273, 1, **1971**, 1194

Y. Ihara, K. Ribet et J.P. Serre, éd., *Galois groups over Q*, Math. Sc. Research Institute Pub. 16, Springer, **1989**

J. Itard, "Les méthodes de Fermat en théorie des nombres", *Revue d'Histoire des Sciences*, III, **1949**, 21 sq.

N. Koblitz, *Introduction to Elliptic Curves and Modular Forms* , Springer, New York Berlin Heidelberg Tokyo, **1984**

S. Lang, *Introduction to Modular Forms*, Springer, New York Berlin Heidelberg, **1976**

B. Mazur, "Number Theory as Gadfly", *Americ. Math. Monthly*, 98, **1991**, 593-610

B. Mazur et H.P.F. Swinnerton-Dyer, "Arithmetic of Weil curves ", *Inventiones Math.*, 25, **1974**, 1-61

J. Morse, *The Reception of Diophantus' "Arithmetic" in the Renaissance*, PhD Princeton, **1981**

O. Neumann, "Uber die Anstösse zu Kummers Schöpfung der idealen complexen Zahlen ", *Mathematical Perspectives*, J. Dauben éd., Academic Press, New York etc., **1981** , 179-200.

R. Noguès, *Théorème de Fermat. Son histoire*, Blanchard, Paris, **1966**

J. Œsterlé, "Nouvelles approches du théorème de Fermat ", *Séminaire Bourbaki*, 40, **1987-1988**, 694.

R. Rashed, *Entre arithmétique et algèbre*, Les Belles Lettres, Paris, **1984**

M. Reid, *Undergraduate algebraic geometry*, London Mathematical Soc. Stud. Texts, 12, Cambridge University Press, Cambridge, **1988**

E. Reyssat, *Quelques aspects des surfaces de Riemann*, Birkhäuser, Bale, **1989**

P. Ribenboim, *13 Lectures on Fermat's Last Theorem*,Springer , New York Heidelberg Berlin, **1979**

K. Ribet, "On modular representations of $Gal(\bar{Q}/Q)$ arising from modular forms", *Inventiones math.*, 100, **1990**, 431-476

K. Ribet, "From the Taniyama-Shimura conjecture to Fermat's Last theorem", *Annales Fac. Sc. Toulouse*, Math., 11, **1990**, 116-139

K. Ribet, "Wiles proves Taniyama conjecture" (sic), *Notices of the AMS*, 40, 6, **1993**, 575-576

M.R. Schroder, *Number Theory in Science and Communication*, Springer, New York Berlin Heidelberg, **1984**

J. P. Serre, "Sur les représentations modulaires de degré 2 de $Gal(\bar{Q}/Q)$ ", *Duke Mathematical Journal*, 54, **1987**, 179-230

G. Shimura, *Introduction to the Arithmetic Theory of Automorphic functions*, Iwanami Shoten et Princeton University Press , **1971**

J. Silverman, *The Arithmetic of elliptic curves*,Springer , Berlin New York Heidelberg Tokyo, **1986**

I. Stewart, *Galois Theory*, Chapman et Hall, London, **1973**

L. Washington, *Introduction to Cyclotomic Fields*, Springer, Berlin New York Heidelberg, **1982**

A. Weil, *Œuvres Scientifiques*, Springer, **1978**

A. Weil, *Number Theory : An Approach through history*, Birkhäuser, Bale, **1983**

D. Zagier, " Modular points, modular curves, modular surfaces and modular forms", *Arbeitstagung Bonn - 1984*, LNM 1111, Springer, **1985**, 225-248

VARIA OPERA
MATHEMATICA
D. PETRI DE FERMAT,
SENATORIS TOLOSANI.

Acceſſerunt ſelectæ quædam ejuſdem Epiſtolæ, vel ad ipſum à pleriſque doctiſsimis viris Gallicè, Latinè, vel Italicè, de rebus ad Mathematicas diſciplinas, aut Phyſicam pertinentibus ſcriptæ.

TOLOSÆ,

Apud JOANNEM PECH, Comitiorum Fuxenſium Typographum, juxta Collegium PP. Societatis JESU.

M. DC. LXXIX.