Les courbes elliptiques; théorème de Mordell-Weil

Bruno Winckler
sous la direction de Marc Hindry

« À l'illustre beauté d'une lune ecliptique,
Rien ne peut l'égaler né de l'arithmétique,
À cette exception: une courbe elliptique. »
– Sergueï Bernikov

Table des matières

1. Introduction

L'objectif de ce stage de recherche d'environ deux mois était de faire mes armes dans la théorie des courbes elliptiques. Plutôt que de me concentrer sur les problèmes de factorisation, j'ai préféré m'orienter vers le théorème de Mordell-Weil, faisant le choix parmi les suggestions de M. Hindry. Si, à la base, une mise à niveau en géométrie projective me semblait nécessaire et suffisante, il s'est avéré que des prérequis d'algèbre générale, puis essentiellement de théorie algébrique des nombres (qui n'est pas toujours algébrique d'ailleurs !) s'avéraient indispensable pour faire mieux que servilement comprendre les propositions, théorèmes et les assertions logiques qui les justifiaient ; mon objectif était en effet de pouvoir comprendre « pourquoi » les définitions et démonstrations sont avancées telles qu'elles le sont, d'éviter de présenter des raisonnements elliptiques [*] et si possible apporter une touche personnelle, chose difficile si on survole superficiellement les théories en jeu. L'introduction des théories utilisées fait l'objet de la première partie de mon mémoire.

Ensuite, pour travailler sur les courbes elliptiques en elles-mêmes, j'ai principalement suivi le fil du cours sur la théorie que propose M. Hindry dans son livre d'arithmétique, en complétant ses idées dans d'autres livres. L'idée de voir une équation diophantienne comme étant une courbe dans le plan, et considérer les cordes et tangentes s'avère vite fructueux, puisque ces opérations définissent une loi de groupe sur les points rationnels de la courbe définie par l'équation diophantienne ; ceci constitue le début de la seconde partie du mémoire, et justifie les prérequis de géométrie projective. Je travaille ensuite sur les hauteurs, concept qui ne posera pas de problème de conception à ce stade du mémoire, et qui nous occupera un petit moment. Enfin, le théorème de Mordell-Weil (le groupe des points rationnels a un nombre fini de générateurs !) vient conclure la seconde partie, en faisant appel à plusieurs techniques apprises pendant le stage.

Je me suis ensuite évertué à travailler sur la question qui succède logiquement au théorème de Mordell-Weil (toujours sur les conseils de mon maître de stage), à savoir « combien y a-t-il de générateurs ? Peut-on les trouver ? », question qui mène naturellement à la conjecture de Birch et Swinnerton-Dyer. Et pour explorer d'autres aspects

∗. Bien que le sujet le permette... Hum. Pour ma défense, ces jeux de mots vaseux me sont inspirés par l'œuvre de référence [**Win**].

des courbes elliptiques qui semblent se traduire en plusieurs langues mathématiques, je me suis intéressé aux fonctions L des courbes elliptiques, et aux formes modulaires dont le lien avec les courbes elliptiques a été prouvé que très récemment (1994, partiellement, avec le théorème de Fermat, puis 1999). Comme je n'ai pas pu consacrer le temps mérité à ces branches des mathématiques, et que j'ai préféré accentuer mes connaissances bien appuyées plutôt que les recherches récentes et autres questions ouvertes, la dernière partie est très concentrée, et contient aucune démonstration.

Enfin, je remercie M. Hindry pour sa disponibilité et ses réponses patientes et éclairantes, ainsi que mes compagnons d'infortune à Chevaleret, pour m'avoir rejoint dans la lutte contre le sommeil, lors des après-midi de grande chaleur et de digestion.

2. Notions essentielles

2.1. En algèbre générale. — Je travaille uniquement sur des anneaux commutatifs intègres.

***Définition 2.1* (Module).** — Un A-module est à l'anneau A ce qu'un K-espace vectoriel est au corps K.

Toutes les définitions et propositions de base à ce sujet (je pense par exemple aux théorèmes de structure) se trouvent dans tout cours d'algèbre digne de ce nom, [**Hd1**] par exemple.

Par contre, pour les besoins de la cause je vais introduire un type particulier d'anneau, qui ressemble beaucoup aux corps de fractions. Un anneau *local* est un anneau qui n'a qu'un seul idéal maximal. C'est une structure suffisamment simple pour qu'on n'ait pas besoin d'être convaincu qu'elle revêt un intérêt particulier. Cependant, on se rend vite compte que les anneaux classiques ne vérifient pas ceci (je ne peux que citer l'anneau $K[[X]]$ des séries formelles, d'idéal maximal unique (X))... Mais dans un anneau quelconque, on peut encore *localiser* « en des idéaux premiers $\mathfrak{p}$ ».

Pour ce faire, soit $\mathfrak{p}$ un idéal premier, et $S = A \setminus \mathfrak{p}$ qui est stable par multiplication d'après la définition d'un idéal premier. Je considère la relation d'équivalence sur $A \times S : (a, s) \sim (a', s') \Leftrightarrow as' - a's = 0$ [(*)]. Avec ces notations, j'en arrive à la

***Définition 2.2* (Anneau, module localisés).** — On appelle anneau localisé en $\mathfrak{p}$, qu'on note $A_\mathfrak{p}$, l'ensemble $(A \times S)/\sim$. Les éléments (a, s) de $A_\mathfrak{p}$ sont notés $\frac{a}{s}$. De même, si M est un A-module, les éléments $\frac{m}{s}$ construits de la même manière constituent un A-module, appelé module localisé en $\mathfrak{p}$.

Si par exemple $\mathfrak{p} = (p)$ où p est un élément de A, on peut voir $A_\mathfrak{p}$ comme les fractions au dénominateur non divisible par p. Si bien qu'en prenant $\mathfrak{p} = \{0\}$, on retrouve le corps des fractions classique. Il est facile de vérifier que c'est un anneau qui porte bien son nom, puisqu'il a pour unique idéal maximal l'ensemble des fractions $\frac{a}{s}$ où $a \in \mathfrak{p}$ (ce qui rend la fraction non inversible dans $A_\mathfrak{p}$; le quotient donne effectivement un corps). On peut aussi considérer le complémentaire d'un ensemble

∗. Dans le cas non intègre, remplacer par « $\exists s_0; s_0(as' - a's) = 0$ », sinon la transitivité pose problème.

d'idéaux premiers pour S, qui est encore multiplicatif. Un intérêt, et qui n'est pas des moindres, dans le processus de localisation est de rendre principal dans l'anneau localisé un idéal qui ne l'était pas dans l'anneau d'origine, pourvu qu'il soit le produit des idéaux premiers par rapport auxquels on localise (ce qui sera le cas dans mon étude, comme je le montrerai dans quelques pages).

Les anneaux locaux peuvent heureusement être reliés à l'anneau tout entier (de même pour les modules), par la proposition suivante qui n'est pas sans rappeler les décompositions en éléments simples.

Proposition 2.3 **(Décomposition en modules localisés)**
Si M est un A-module de longueur finie[†]*, alors on a la décomposition*

$$M \simeq \bigoplus_{\mathfrak{p}} M_{\mathfrak{p}}.$$

La proposition subsiste pour un anneau commutatif artinien[‡].

Démonstration. — La démonstration se fait par récurrence sur la longueur l. Si $l = 1$, M est de la forme $A/\mathfrak{M}$ qui est un corps, et cette propriété permet de voir que $(A/\mathfrak{M})_{\mathfrak{M}} = A/\mathfrak{M}$ et $(A/\mathfrak{M})_{\mathfrak{p}} = 0$ si $\mathfrak{p} \neq \mathfrak{M}$. Ensuite, si le résultat est connu pour M_1 (de longueur $\leqslant l-1$) et M/M_1 qui est simple, j'en déduis le résultat pour M. □

Enfin, un résultat qui n'est pas lié aux modules, mais aux zéros des polynômes... Ce qui a son importance, puisqu'en géométrie projective les objets sont essentiellement définis par des équations polynomiales. Ce théorème des zéros de Hilbert, ou *Nullstellensatz*, servira quand je parlerai des hauteurs, et semble être également le fer de lance de la géométrie algébrique.

Théorème 2.4 **(*Nullstellensatz* de Hilbert)**. — *Soient $P_1, \ldots, P_m$ des polynômes dans $K[X_1, \ldots, X_m]$ avec K algébriquement clos. Si Q est un polynôme s'annulant sur le lieu des zéros communs des P_i, alors une puissance de Q est dans l'idéal engendré par les P_i.*

La démonstration est en annexe.

2.2. En théorie algébrique des nombres. — Dans cette partie, un anneau A intègre sera toujours suivi de près par K qui désignera son corps de fractions. Si le contexte fait plutôt parler de $\mathbb{Z}$ ou $\mathbb{Q}$, alors K désignera une extension finie de $\mathbb{Q}$ (qu'on appelle un *corps de nombres*)... Qui sera également le corps de fractions de $\mathcal{O}_K$, anneau ici introduit !

†. Je rappelle qu'un A-module est de longueur finie s'il existe une suite de sous-modules $M = M_0 \supset M_1 \supset \cdots \supset M_l = \{0\}$ telle que chaque M_i/M_{i+1} soit de la forme $A/\mathfrak{M}$ avec $\mathfrak{M}$ idéal maximal. Le théorème de Jordan-Hölder assure l'unicité à permutation près d'une telle suite, donc la longueur l est unique. Voir [**Lng**].

‡. Un anneau artinien est un anneau où toute suite *décroissante* d'idéaux est stationnaire. C'est un peu le pendant d'un noethérien, et un anneau à la fois artinien et noethérien est de longueur finie. Voir [**Lng**].

***Définition 2.5* (Entier algébrique).** — Si $A \subseteq B$, on dit que $x \in B$ est un entier de B sur A s'il existe un polynôme unitaire à coefficients dans A qui annule x.

Ainsi, par exemple, $\sqrt{2}$ est un entier sur $\mathbb{Z}$, car $X^2 - 2$ l'annule. De même, $\frac{\sqrt{5}+1}{2}$ est un entier de $\mathbb{Q}(\sqrt{5})$ sur $\mathbb{Z}$, puisqu'il est annulé par $X^2 - X - 1$, mais il n'est pas entier de $\mathbb{Q}$ sur $\mathbb{Z}$. Les éléments entiers de K sur A joueront un rôle particulier.

Proposition 2.6. — *Un élément de K est entier sur A si et seulement si l'anneau engendré par cet élément est contenu dans un A-module de type fini (ou est lui-même de type fini).*

Corollaire 2.7. — *L'ensemble des éléments de K entiers sur A forme un anneau.*

Pour démontrer le corollaire, il suffit de remarquer que $A[\alpha + \beta]$ et $A[\alpha\beta]$ sont contenus dans $A[\alpha, \beta]$.

Corollaire 2.8. — *Soit $A \subseteq B$. Si α est entier sur B, et si chaque élément de B est entier sur A, alors α est entier sur A.*

Là encore, il n'y a pas trop de difficulté pour s'en convaincre.

Démonstration de la proposition 2.6. — Si α est entier sur A, on peut écrire une relation du type $\alpha^n = \sum_{i=0}^{n-1} a_i \alpha^i$ avec les a_i dans A. Alors, $A[\alpha]$ est de type fini car égal à $A + A\alpha + \cdots + A\alpha^{n-1}$. Réciproquement, si $A[\alpha]$ est dans $Au_1 + \cdots + Au_n$, on peut écrire chaque αu_i sous la forme $\sum_{j=1}^{m} a_{ij} u_j$. Alors, $P = \det(Xid - M)$ où $M = ((a_{ij}))_{i,j}$ annule α (on utilise le théorème de Cayley-Hamilton !), et est bien unitaire à coefficients dans A. □

Dorénavant, on note $\mathscr{O}_K$ cet anneau, et on l'appelle *clôture intégrale* de A. Si A égale sa clôture intégrale, on dit que A est *intégralement clos*. Il est facile de voir qu'un anneau factoriel (en particulier un anneau principal) est intégralement clos, en suivant la même démonstration que dans le cas de $\mathbb{Z}$ [*]. Pour $\mathbb{Q} \subseteq K$ une extension finie, l'anneau $\mathscr{O}_K$ a lui aussi le bon goût d'être intégralement clos : son corps de fractions est K (remarquer que si α est algébrique sur $\mathbb{Q}$, il existe un entier d tel que $d\alpha$ soit un entier algébrique), et un élément de K entier sur $\mathscr{O}_K$ l'est aussi sur $\mathbb{Z}$ d'après le corollaire précédent. Il est donc dans $\mathscr{O}_K$.

L'anneau des entiers a le défaut de ne pas être factoriel. Le cas historique le plus fâcheux faisant intervenir des anneaux non factoriels est celui concernant l'équation de Fermat $x^p + y^p = z^p$ avec p premier [†]. Heureusement, pour une certaine classe

∗. Si $x = \frac{p}{q} \in \mathbb{Q}$ avec p et q premiers entre eux est annulé par un polynôme unitaire à coefficients entiers, on peut écrire une relation du type $(p/q)^n + \sum_{i=0}^{n-1} a_i (p/q)^i = 0$, donc $p^n + \sum_{i=0}^{n-1} a_i p^i q^{n-i} = 0$. Dans $\mathbb{Z}/q\mathbb{Z}$, on a alors $p^n \equiv 0 \mod q$, donc q divise p et $q \in \{\pm 1\}$, donc $x = \pm p \in \mathbb{Z}$. La factorialité est nécessaire pour utiliser le théorème de Gauss.

†. Elle fait intervenir l'anneau des entiers de $\mathbb{Q}(\exp(2i\pi/p))$.

d'anneaux décrite ci-dessous, une autre propriété remplace en quelque sorte la factorialité : c'est la factorisation des idéaux. Dans mon étude, seul le cas de $\mathscr{O}_K$ nous intéresse, et il est plus facile dans ce cas particulier d'établir le résultat. Mais par satisfaction intellectuelle, j'ai préféré traiter le cas général, ce qui m'a obligé à établir avant tout quelques résultats sur les idéaux dans un anneau noethérien. Comme par exemple les propositions suivantes :

Proposition 2.9. — *Si un idéal premier I contient un produit d'idéaux $\mathfrak{p}_1 \cdots \mathfrak{p}_n$, alors il contient l'un d'eux.*

Je rappelle que le produit IJ d'idéaux désigne l'ensemble des $\sum x_i y_i$ où $x_i \in I$ et $y_i \in J$, et non pas l'ensemble des $x_i y_i$ qui n'a pas de raison d'être idéal.

Démonstration. — *Reductio ad absurdum*, si ce n'est pas le cas, pour tout i il existe $a_i \in \mathfrak{p}$ tel que $a_i \not\equiv 0 \mod I$. Alors, $a_1 \cdots a_n \not\equiv 0 \mod I$ par intégrité de A/I, ce qui est impossible car $a_1 \cdots a_n \in I$! □

Proposition 2.10. — *Dans un anneau noethérien, tout idéal non nul contient un produit d'idéaux premiers non nuls.*

Démonstration. — *Reductio ad absurdum*, je suppose que l'ensemble des idéaux non nuls ne vérifiant pas la propriété voulue est non vide [‡]. Cet ensemble Φ a alors un élément maximal (au sens de l'inclusion), noté I. L'idéal I n'est pas premier, car sinon $I \subseteq I$ et $I \notin \Phi$. Il existe donc x et y dans $A \smallsetminus I$ tels que $xy \in I$. Ainsi $I + Ax$ et $I + Ay$ sont des idéaux qui contiennent strictement I, donc ne sont pas dans Φ et ils contiennent un produit d'idéaux. Donc $(I + Ax)(I + Ay)$ également, or $(I + Ax)(I + Ay) = I + Ix + Iy + xyA \subseteq I$ donc I en contient un ! □

Anneaux de Dedekind et idéaux fractionnaires. — Pour poursuivre, il est indispensable de savoir qu'un *idéal fractionnaire* est un sous-A-module de K tel que $dI \subseteq A$ pour un certain $d \in K^*$ (appelé dénominateur commun). Pour $d = 1$, on retrouve un idéal classique, parfois appelé *entier*.

Définition 2.11 **(Anneau de Dedekind)**. — Un anneau de Dedekind est un anneau intègre noethérien, intégralement clos, dont tout idéal premier non nul est maximal.

Par exemple, les $\mathscr{O}_K$ sont des anneaux de Dedekind. Les propriétés se vérifient à l'aide du constat que le quotient de $\mathscr{O}_K$ par un idéal non nul est fini (voir plus tard la proposition 2.16, et en se rappelant qu'un anneau fini est intègre si et seulement si c'est un corps). Plus généralement, un anneau principal est de Dedekind, $K[T]$ par exemple.

Proposition 2.12. — *Tout idéal premier d'un anneau de Dedekind est inversible dans le monoïde des idéaux fractionnaires.*

‡. Procéder ainsi pour profiter du caractère noethérien et conclure par l'absurde, avec l'idéal maximal, est souvent très efficace !

Démonstration. — Soit I l'idéal premier non nul (donc maximal) en question. Je cherche un idéal fractionnaire J tel que $IJ = A$. À cet effet, il paraît juste de définir J comme étant $\{x \in K | xI \subseteq A\}$ (qui est bien un idéal fractionnaire, de dénominateur commun tout élément de K^*). Le lecteur en exercice vérifiera que la définition de J fournit les inclusions $I \subseteq IJ \subseteq A$. Alors, par maximalité de I, on a $IJ \in \{I, A\}$. Je n'ai plus qu'à montrer que $IJ = I$ est impossible ; ceci conduirait à $J = A$: l'anneau A étant intégralement clos, $IJ = I$ signifie que pour tout $x \in J$ et tout n, $x^n I \subseteq I$, donc $A[x]$ est un idéal fractionnaire de A donc est de type fini (A est noethérien !), donc x est entier. Or je suis capable d'exhiber un élément de J qui n'est pas dans A.

En effet, soit $a \in I$ non nul. L'idéal $aA \subseteq I$ est non nul, donc contient un produit d'idéaux premiers $\mathfrak{p}_1 \cdots \mathfrak{p}_n$, et je prends n minimal tant qu'à faire. Alors I contient l'un des $\mathfrak{p}_i$, par exemple $\mathfrak{p}_1$, parce qu'il est premier. En gardant à l'esprit que les idéaux premiers non nuls sont ici maximaux, je déduis $I = \mathfrak{p}_1$. Si je note $\mathfrak{B} = \mathfrak{p}_2 \cdots \mathfrak{p}_n$, alors $I\mathfrak{B} \subseteq aA$ et $\mathfrak{B} \not\subseteq aA$ par minimalité de n. Bref, il existe $b \in \mathfrak{B}$ tel que $b \notin aA$, et alors $ba^{-1} \notin A$, tandis que $Iba^{-1} \subseteq A$ (donc $ba^{-1} \in J$). D'où $IJ = A$. □

Je note I^{-1} l'inverse de I dorénavant.

Théorème 2.13. — *Soit A un anneau de Dedekind. Tout idéal fractionnaire $\mathfrak{M}$ peut être écrit de manière unique sous la forme*

$$\mathfrak{M} = \prod_{\mathfrak{p}} \mathfrak{p}^{\mathrm{ord}_{\mathfrak{p}}(\mathfrak{M})}$$

où les $\mathrm{ord}_{\mathfrak{p}}(\mathfrak{M})$ *sont presque tous nuls (tous, sauf un nombre fini d'entre eux).*

C'est en particulier le cas pour les anneaux $\mathscr{O}_K$, ce qui sera utile au moment de définir les hauteurs.

Corollaire 2.14. — *L'ensemble des idéaux fractionnaires de ce même A forme un groupe, d'élément neutre A.*

Le corollaire se déduit aisément du théorème, en prenant les inverses de chaque idéal premier.

Démonstration. — Si $\mathfrak{M}$ est un idéal fractionnaire de dénominateur commun d, alors $\mathfrak{M} = (d\mathfrak{M})(dA)^{-1}$, et j'ai juste à démontrer le résultat pour les idéaux entiers. *Reductio ad absurdum*, si l'ensemble Φ des idéaux non nuls qui ne vérifient pas la propriété voulue est non vide, soit I son idéal maximal (au sens de l'inclusion). En particulier, I est différent de A, qui est égal au produit vide des idéaux premiers. Soit $\mathfrak{p}$ l'idéal maximal (pour de vrai cette fois) contenant I. Alors, $I \subseteq \mathfrak{p} \subsetneq A$, donc $I\mathfrak{p}^{-1} \subseteq \mathfrak{p}\mathfrak{p}^{-1} = A$. Par un raisonnement proche de celui de la proposition 2.12, je montre que $I \subsetneq I\mathfrak{p}^{-1}$, et par conséquent $I\mathfrak{p}^{-1} \notin \Phi$, donc admet une décomposition en facteurs premiers. En multipliant par $\mathfrak{p}$, I a aussi une telle décomposition, ce qui est absurde.

L'unicité découle du fait que si $\prod_{i=1}^{r} \mathfrak{p}_i^{\alpha_i} = \prod_{i=1}^{s} \mathfrak{q}_i^{\beta_i}$ (où je supprime les doublons et regroupe les puissances de deux éventuelles décompositions, si bien que tous les exposants sont positifs), alors $\prod_{i=1}^{s} \mathfrak{q}_i^{\beta_i} \subseteq \mathfrak{p}_1$, donc $\mathfrak{p}_1$ contient l'un d'eux, et par maximalité il y a égalité, or j'ai tout regroupé... □

On peut étudier de plus près les $\mathrm{ord}_{\mathfrak{p}}(\mathfrak{M})$. Il est très simple de montrer que $\mathfrak{a} \subseteq \mathfrak{b}$ implique $\mathrm{ord}_{\mathfrak{p}}(\mathfrak{b}) \leqslant \mathrm{ord}_{\mathfrak{p}}(\mathfrak{a})$, que $\mathrm{ord}_{\mathfrak{p}}(\mathfrak{ab}) = \mathrm{ord}_{\mathfrak{p}}(\mathfrak{a}) + \mathrm{ord}_{\mathfrak{p}}(\mathfrak{b})$, ou encore que $\mathfrak{b} \subseteq A \Leftrightarrow \mathrm{ord}_{\mathfrak{p}}(\mathfrak{b}) \geqslant 0$ pour tout $\mathfrak{p}$. Ainsi, $\mathscr{O}_K$ peut être vu comme l'ensemble $\{x \in K \mid |\forall \mathfrak{p}, \mathrm{ord}_{\mathfrak{p}}(x) \geqslant 0\}$ où $\mathrm{ord}_{\mathfrak{p}}(x)$ est bien sûr $\mathrm{ord}_{\mathfrak{p}}(x\mathscr{O}_K)$.

Géométrie des nombres. — Pour avoir une bonne compréhension de mon sujet d'étude, il était nécessaire de connaître tous les résultats de base sur $\mathscr{O}_K$, nouvel ensemble d'un intérêt arithmétique incontestable. Je me suis encore attelé à montrer que pour un corps de nombres K, le groupe des classes d'idéaux $\mathcal{CL}(K)$ (pour la relation d'équivalence $I \sim J \Leftrightarrow IJ^{-1}$ est fractionnaire) est fini. Une notion de « taille » peut être associée à un idéal (un nombre entier pour être précis), et pour montrer la finitude des groupes des classes d'idéaux on peut montrer que la « taille » d'un idéal est bornée, et enfin qu'il n'y a qu'un nombre fini d'idéaux ayant une certaine « taille ». Ceci m'a conduit à découvrir un autre aspect de l'arithmétique, à savoir la géométrie des nombres dont un résultat de base est le théorème de Minkowski.

Pour commencer, je définis la norme N (dans K) d'un élément α de K : il s'agit du déterminant de la multiplication par α, vue comme application $\mathbb{Q}$-linéaire de K dans K. Il n'est pas inutile de définir la trace (dans K) de α comme étant la trace (!) de cette même application. Si on connait le polynôme minimal de α, on a une expression assez concrète de ces applications :

Proposition 2.15. — *Si α est algébrique sur $\mathbb{Q}$ et $K = \mathbb{Q}(\alpha)$, alors*

$$\mathrm{N}(\alpha) = \prod_{i=1}^{d} \alpha_i, \quad \mathrm{tr}(\alpha) = \sum_{i=1}^{d} \alpha_i, \tag{1}$$

où les α_i sont les racines du polynôme minimal de α sur $\mathbb{Q}$. Si $\alpha \in K$, alors

$$\mathrm{N}(\alpha) = \left(\prod_{i=1}^{d} \alpha_i\right)^{[K:\mathbb{Q}(\alpha)]}, \quad \mathrm{tr}(\alpha) = [K : \mathbb{Q}(\alpha)] \sum_{i=1}^{d} \alpha_i.$$

Démonstration. — C'est immédiat dans le cas $K = \mathbb{Q}(\alpha)$ en remarquant que le polynôme caractéristique de la multiplication par α est le polynôme minimal de α : si on choisit comme base les α^i, la matrice de l'application est une matrice compagnon, ce qui facilite la tâche ! Dans le cas $m = [K : \mathbb{Q}(\alpha)]$, il suffit de remarquer que le polynôme caractéristique est, cette fois, le polynôme minimal à la puissance m... □

Cette proposition montre que si $\alpha \in \mathscr{O}_K$, alors sa norme et sa trace sont des entiers algébriques et rationnels, donc des entiers. Cette proposition peut être utilisée pour déterminer l'anneau des entiers dans certains cas (extension quadratique, ou

engendrée par une racine de l'unité par exemple). On peut aussi exprimer ces quantités autrement, plus pratique selon les situations : d'après le théorème de l'élément primitif pour l'extension de K sur $\mathbb{Q}$, il existe $\gamma \in K$ tel que $K = \mathbb{Q}(\gamma)$. Le polynôme minimal sur $\mathbb{Q}$ de γ a d'autres racines, r_1 racines réelles (notées γ_i pour $i \in [\![1, r_1]\!]$) et r_2 *paires* de racines complexes (notées $(\gamma_i, \bar{\gamma_i})$ pour $i \in [\![r_1+1, r_1+r_2]\!]$), où $r_1 + 2r_2 = n$. Définissons alors les plongements de K dans $\mathbb{C}$: les r_1 plongements *réels* σ_i sont des morphismes $K \to \mathbb{C}$ définis par $\gamma \mapsto \gamma_i$. Les $2r_2$ plongements *complexes* σ_i, $\bar{\sigma_i}$ sont définis par $\sigma_i : \gamma \mapsto \gamma_i$, $\bar{\sigma}_i : \gamma \mapsto \bar{\gamma_i}$. Comme ce sont des morphismes de corps, l'ensemble $\{x \in K \mid f(x) = x\}$ est un sous-corps de K donc contient $\mathbb{Q}$, si bien que les σ_i sont l'identité sur $\mathbb{Q}$, et définir l'image de γ suffit à les définir entièrement. Alors, une fois ceci acquis, on vérifie très vite que

$$\mathrm{N}(\alpha) = \prod_{i=1}^{r_1} \sigma_i(\alpha) \prod_{i=r_1+1}^{r_1+r_2} |\sigma_i(\alpha)|^2, \quad \mathrm{tr}(\alpha) = \sum_{i=1}^{r_1} \sigma_i(\alpha) + 2 \sum_{i=r_1+1}^{r_1+r_2} \Re(\sigma_i(\alpha)) \qquad (2)$$

On garde les mêmes définitions pour r_1 et r_2, parce qu'elles nous serviront encore à plusieurs reprises.

On étend la définition de la norme aux idéaux principaux facilement, en posant $\mathrm{N}(\alpha\mathscr{O}_K) = |\mathrm{N}(\alpha)|$. Pour un idéal quelconque non nul, on pose $\mathrm{N}(I) = \mathrm{card}\,(\mathscr{O}_K/I)$. Je vérifie d'abord que $\mathrm{N}(I)$ est bien défini (ou plutôt fini, en économisant une syllabe), puis que dans le cas d'un idéal principal, les deux définitions coïncident. Ce qui assure que la norme est finie est :

Proposition 2.16. — *Si $[K : \mathbb{Q}] = n$, alors $\mathscr{O}_K$ est un $\mathbb{Z}$-module de rang n. Plus généralement, c'est le cas pour un idéal non nul de $\mathscr{O}_K$.*

Dans tel cas, la norme est bien finie, puisqu'alors le quotient serait isomorphe à un anneau du type $\prod_{i=1}^{n} \mathbb{Z}/a_i\mathbb{Z}$, de cardinal fini.

Démonstration. — Tout d'abord, $\mathscr{O}_K$ est de rang au moins n, car si $(f_i)_{1 \leqslant i \leqslant n}$ est une $\mathbb{Q}$-base de K, quitte à les multiplier par un dénominateur commun je peux les supposer entiers algébriques, et alors $\mathbb{Z}f_1 + \cdots + \mathbb{Z}f_n \subseteq \mathscr{O}_K$. Montrer que le rang est exactement n nécessite un peu plus d'astuce... Comme la trace : $K \times K \to \mathbb{Q}$ est non dégénérée (par exemple parce que $\mathrm{tr}(xx^{-1}) = \mathrm{tr}(1) = [K : \mathbb{Q}] \neq 0$), il existe des $(f_j^*)_{1 \leqslant i \leqslant n}$ tels que $\mathrm{tr}(f_i f_j^*) = \delta_{ij}$ (symbole de Kronecker). [§] Alors, si je note d un entier tel que les df_j^* soient entiers algébriques, pour tout x entier algébrique j'ai $\mathrm{tr}(dxf_j^*) = dx_i \in \mathbb{Z}$, donc $\mathscr{O}_K \subseteq \frac{1}{d}(\mathbb{Z}f_1 + \cdots \mathbb{Z}f_n)$, et $\mathscr{O}_K$ est de rang au plus n. □

Pour un idéal I, je considère α tel que $\alpha\mathscr{O}_K \subseteq I \subseteq \frac{\alpha}{d}\mathscr{O}_K$. Enfin, je montre que $\mathrm{N}(\alpha\mathscr{O}_K) = |\mathrm{N}(\alpha)|$ en remarquant que $\alpha\mathscr{O}_K = \mathrm{Im}(m_\alpha)$ (multiplication par α) et que

§. D'une part, pour toute base $(f_i)_{1 \leqslant i \leqslant n}$ de K, il existe une base duale $(f_j^*)_{1 \leqslant j \leqslant n}$ de $\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{Q})$ (espace dual du $\mathbb{Q}$-espace vectoriel K), qui vérifie par définition la relation $f_j^*(f_i) = \delta_{ij}$. D'autre part, comme la trace est non dégénérée, l'application $K \to \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{Q})$ définie par $x \mapsto (y \mapsto \mathrm{tr}(yx))$ est injective, donc bijective en comparant les dimensions. Il existe donc une unique famille $(g_j)_{1 \leqslant j \leqslant n}$ d'éléments dans K tels que $\mathrm{tr}(\cdot g_j) = f_j^*$ pour tout j, d'où la propriété énoncée ; pour ne pas encombrer les notations, on a identifié g_j et f_j^* dans la démonstration.

très généralement, si $f : \mathbb{Z}^n \to \mathbb{Z}^n$ est une application $\mathbb{Z}$-linéaire de déterminant non nul, alors $\mathbb{Z}^n/f(\mathbb{Z}^n)$ est de cardinal $|\det(f)|$ (se démontre tout seul en utilisant la réduction en matrices diagonales $\mathrm{diag}(a_1, \ldots, a_n)$ où $a_i | a_{i+1}$). Bref, j'ai bien assuré mes arrières.

Une propriété de base très utile de la norme est la suivante :

***Proposition 2.17* (Multiplicativité de la norme).** — *Pour tous idéaux I et J, on a $\mathrm{N}(IJ) = \mathrm{N}(I)\mathrm{N}(J)$.*

Démonstration. — La décomposition en idéaux premiers montre que j'ai juste à démontrer le résultat dans le cas où J est premier non nul (donc maximal). On a d'une part la suite exacte suivante :

$$0 \longrightarrow I/IJ \longrightarrow \mathscr{O}_K/IJ \longrightarrow \mathscr{O}_K/I \longrightarrow 0,$$

qui assure $\mathrm{N}(IJ) = \mathrm{N}(I)\mathrm{card}\ (I/IJ)$, et d'autre part que $\mathbb{F}_{\mathrm{N}(J)} = \mathscr{O}_K/J$ est un corps. Alors, I/IJ est un $\mathbb{F}_{\mathrm{N}(J)}$-espace vectoriel, et s'il est de dimension 1 on a par conséquent $\mathrm{card}\ (I/IJ) = \mathrm{N}(J)$, et le résultat voulu est démontré. Je montre qu'il n'y a pas de sous-espace strict non réduit au vecteur nul. Si $\{0\} \subseteq L \subsetneq I/IJ$, alors L est aussi un sous-module de I/IJ, donc il existe un idéal $IJ \subsetneq I' \subseteq I$ tel que $L = I'/IJ$. Comme $0 \leqslant \mathrm{ord}_J(I') - \mathrm{ord}_J(I) < 1$, et $\mathrm{ord}_{\mathfrak{p}}(I') - \mathrm{ord}_{\mathfrak{p}}(I) = 0$ sinon, $I' = I$, donc $L = \{0\}$. □

On peut enfin commencer à faire de la géométrie des nombres sereinement. Je rappelle qu'un sous-groupe discret Λ de $\mathbb{R}^n$ est isomorphe à un certain $\mathbb{Z}^r$. Si $r = n$, ce groupe est appelé *réseau*. On appelle déterminant (ou volume) de ce réseau le déterminant d'un système générateur du réseau dans la base canonique. Ceci correspond également à la mesure (de Lebesgue) de l'enveloppe convexe formée des vecteurs générateurs.

***Théorème 2.18* (Théorème de Minkowski).** — *Soit $C \subseteq \mathbb{R}^n$ une partie compacte, convexe et symétrique par rapport à $\vec{0}$, et Λ un réseau. Si $\lambda_n(C) \geqslant 2^n \det(\Lambda)$, alors $C \cap \Lambda$ contient un vecteur non nul.*

Je l'utiliserai à plusieurs reprises par la suite, mais pour la beauté du geste je propose une application presque immédiate de ce résultat, qui illustre le mariage entre arithmétique et géométrie :

***Corollaire 2.19*.** — *Si p est un nombre premier congru à 1 modulo 4, alors p est somme de deux carrés entiers.*

Puisqu'il est ici question de géométrie, on peut interpréter cette proposition ainsi : si p est un nombre premier congru à 1 modulo 4, alors il existe $(x, y) \in \mathbb{Z}^2$ tel que $\|(x, y)\|_2 = \sqrt{p}$. Ce qui donne une idée du compact à considérer pour la démonstration (qui est en annexe).

De même, on sait montrer qu'un entier est somme de quatre carrés. Mais revenons à la démonstration du théorème de Minkowski :

Démonstration. — Un changement de base linéaire conduit facilement au cas $\Lambda = \mathbb{Z}^n$. Remarquons que pour $T \subseteq \mathbb{R}^n$ quelconque :

$$\lambda_n(T) = \sum_{\vec{v}\in\mathbb{Z}^n} \lambda_n(T \cap ([0,1[^n + \vec{v})) = \sum_{\vec{v}\in\mathbb{Z}^n} \lambda_n((T-\vec{v}) \cap [0,1[^n).$$

Si je suppose qu'en plus tous les $T - \vec{v}$ sont disjoints, je peux finir le calcul pour trouver $\lambda_n(T) \leqslant \lambda_n([0,1[^n) = 1$. Alors, par contraposée, si $\lambda_n(T) > 1$, je peux trouver $\vec{x}$ dans un certain $T \cap (T + \underbrace{\vec{\mu}}_{\neq\vec{0}})$ et de plus $\vec{\mu} \in T - T$. Il est donc intéressant de voir que si $\lambda_n(C) > 2^n$, alors on peut se ramener « au cas T » via la transformation linéaire $\vec{x} \mapsto \frac{1}{2}\vec{x}$ de déterminant 2^{-n}, et il existe alors $\vec{\mu} \in T - T = C$ à coordonnées entières. Si $\lambda_n(C) = 2^n$, on conclut facilement en considérant les $(1+1/m)C$, et en se rappelant qu'une suite convergente à valeurs dans $\mathbb{Z}^n$ stationne [¶]. □

À présent, notons $\sigma(x) = (\sigma_1(x), \dots, \sigma_{r_1+r_2}(x)) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$, le *plongement canonique* de K dans $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \simeq \mathbb{R}^n$. Il vérifie la proposition (bientôt utile) suivante :

Proposition 2.20. — *Si M est un sous-$\mathbb{Z}$-module libre de rang n de K, et si $(x_i)_{1\leqslant i\leqslant n}$ est une $\mathbb{Z}$-base de M, alors $\sigma(M)$ est un réseau de $\mathbb{R}^n$, dont le volume est donné par* $\det(\sigma(M)) = 2^{-r_2}|\det(\sigma_i(x_j))|$.

On pose $d = \det(\sigma_i(x_j))$ dorénavant, pour alléger les notations. La démonstration s'obtient par un calcul formel, avec les qualités du déterminant (linéarité par rapport aux lignes, caractérisation de l'indépendance de vecteurs). En appliquant ce résultat à $\mathscr{O}_K$ et $\mathfrak{a}$ un idéal, on obtient encore une fois des réseaux, et $\det(\sigma(\mathscr{O}_K)) = 2^{r_2}\sqrt{|d|}$, $\det(\sigma(\mathfrak{a})) = 2^{r_2}\sqrt{|d|}\mathrm{N}(\mathfrak{a})$.

Bref ! Je reviens sur l'objectif annoncé, à savoir borner la « taille » (c'est-à-dire la norme) des idéaux. Un résultat préliminaire indispensable est la

Proposition 2.21. — *L'idéal $\mathfrak{a}$ contient un élément $x \in K^*$ tel que*

$$|\mathrm{N}(x)| \leqslant \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n}|d|^{1/2}\mathrm{N}(\mathfrak{a}) = C \cdot \mathrm{N}(\mathfrak{a}).$$

On a déjà vu à l'instant $\mathrm{N}(\mathfrak{a})|d|^{1/2}2^{-r_2} = \det(\sigma(\mathfrak{a}))$, et on aimerait $\geqslant u|\mathrm{N}(x)|$ avec u une certaine quantité, qui sera celle de la proposition (qui dépend de la démonstration proposée, et qui importe peu en vérité). Ça sent le Minkowski à plein nez...

Démonstration. — On sait que $|\mathrm{N}(x)| = \prod_{i=1}^{r_1} |\sigma_i(x)| \prod_{j=r_1+1}^{r_1+r_2} |\sigma_j(x)|^2$. Si j'utilise l'inégalité arithmético-géométrique pour linéariser sous quelques facettes, j'obtiens

$$|\mathrm{N}(x)| \leqslant \left(\frac{1}{n}\sum_{i=1}^{r_1}|\sigma_i(x)| + \frac{2}{n}\sum_{j=r_1+1}^{r_1+r_2}|\sigma_j(x)|\right)^n.$$

¶. C'est dans le cas $\lambda_n(C) = 2^n$ que la compacité sert.

Si, alors, j'arrive à trouver un élément dans $\sigma(\mathfrak{a})$ et dans

$$B_t = \left\{ (y_1, \dots, y_{r_1}, z_1, \dots, z_{r_2}\} \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \mid \quad \sum_{i=1}^{r_1} |y_i| + 2\sum_{j=1}^{r_2} |z_j| \leqslant t \right\}$$

pour un certain t (il s'agit bien d'un compact symétrique convexe), j'obtiendrai l'inégalité voulue. B_t a un volume calculable par récurrence sur r_1 et r_2. Je trouve $\lambda_n(B_t) = 2^{r_1} \left(\frac{\pi}{2}\right)^{r_2} \frac{t^n}{n!}$, et si je prends t tel que $\lambda_n(B_t) = 2^n \det(\sigma(\mathfrak{a}))$ (c'est-à-dire $t^n = 2^{n-r_1}\pi^{r_2} n! |d|^{1/2} \mathrm{N}(\mathfrak{a}) > 0$), alors le théorème de Minkowski s'applique ! Je peux exhiber l'objet désiré au début de la démonstration, et

$$|\mathrm{N}(x)| \leqslant \frac{t^n}{n!} \leqslant \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n} |d|^{1/2} \mathrm{N}(\mathfrak{a}).$$

□

Alors, comme la norme est multiplicative, on peut écrire $\mathrm{N}(x\mathfrak{a}^{-1}) \leqslant C$. Ceci permet de montrer que toute classe d'idéaux contient un idéal entier $\mathfrak{b}$ inférieur à C : prendre $x\mathfrak{a}^{-1}$ où $\mathfrak{a}$ est un idéal de l'inverse de la classe considérée (il est entier, quitte à le multiplier par un dénominateur commun : il reste dans la même classe). Je peux donc construire une injection de $\mathcal{CL}(K)$ dans Φ l'ensemble des idéaux entiers de norme inférieure à C.

Corollaire 2.22. — *Le groupe des classes d'idéaux est fini.*

Ceci provient de l'injection citée, et du fait que Φ est en fait fini.

Démonstration. — Il suffit de montrer que pour chaque entier q, il n'existe qu'un nombre fini d'idéaux (modulo $\sim$) de norme q. Soit I tel que $\mathrm{card}\ (\mathscr{O}_K/I) = q$. Alors, comme l'ordre d'un élément de $G = \mathscr{O}_K/I$ divise le cardinal de G (théorème de Lagrange), $q \cdot 1 \equiv 0$ dans G, donc $q \in I$ et mieux, $q\mathscr{O}_K \subseteq I$. D'où

$$0 \leqslant \mathrm{ord}_{\mathfrak{p}}(I) \leqslant \mathrm{ord}_{\mathfrak{p}}(q\mathscr{O}_K)$$

pour tout $\mathfrak{p}$, ce qui laisse un choix limité de valeurs possibles pour les $\mathrm{ord}_{\mathfrak{p}}(I)$ et donc pour I...

ÉPIC

□

Je fais remarquer que tous les résultats mentionnés m'ont intéressé non seulement pour leur application aux courbes elliptiques, mais aussi parce qu'elle permet de traiter l'équation de Fermat $x^p + y^p = z^p$ dans un certain nombre de cas, le cas où p est un nombre premier régulier. Pour une démonstration, voir [**B & C**].

Pour conclure la description de $\mathscr{O}_K$ (oui, ce n'est pas fini !), je démontre un théorème *violent* qui servira dans les dernières étapes de la démonstration du théorème de Mordell-Weil :

Théorème 2.23 **(Théorème des unités de Dirichlet)**. — *Avec les mêmes notations, soit $r = r_1 + r_2 - 1$. Alors $(\mathscr{O}_K^*, \cdot)$ est un groupe abélien de type fini. Plus précisément, il est isomorphe à $\mathbb{Z}^r \times \mathbb{U}_K$, où $\mathbb{U}_K$ est le groupe (fini) des racines de l'unité de K.*

La partie la plus difficile de la démonstration est le fait que le rang de $\mathscr{O}_K^*$ (en tant que groupe) est exactement r. Ce n'est pas nécessaire pour la suite, et surtout assez long, donc je renvoie à [**Sam**], et signale qu'on utilise le théorème de Minkowski qui est décidément très utile. Avant de démarrer la démonstration, il est important de remarquer que les entiers inversibles sont exactement les entiers de norme 1 ou -1 : si x est inversible, $x\mathscr{O}_K = \mathscr{O}_K$ et $|N(x)| = 1$. Réciproquement, si x entier vérifie $\mathrm{N}(x) = \pm 1$, alors x vérifie, quitte à multiplier par -1, une relation polynomiale de la forme

$$\pm x^n + a_{n-1}x^{n-1} + \cdots - 1 = 0,$$

donc $\pm(x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_1)x = 1$, donc x est inversible dans $\mathscr{O}_K^*$. Mais revenons à la démonstration du théorème des unités :

Démonstration. — Pour montrer que le groupe est de type fini, la tradition veut qu'on considère le morphisme suivant, appelé *plongement logarithmique* de K^* :

$$\mathscr{L} : \begin{cases} K^* & \to & \mathbb{R}^{r_1+r_2} \\ \alpha & \mapsto & (\ln(|\sigma_1(\alpha)|), \ldots, \ln(|\sigma_{r_1}(\alpha)|), 2\ln(|\sigma_{r_1+1}(\alpha)|), \ldots, 2\ln(|\sigma_{r_1+r_2}(\alpha)|)) \end{cases}$$

et on montre alors que $\{x \in \mathscr{O}_K^* \quad | \quad \mathscr{L}(x) \in B\}$ est fini pour tout compact B. Vous comprendrez bientôt pourquoi ce choix de morphisme est plus naturel qu'il n'y paraît. En tout cas, ce qu'on veut montrer aurait pas mal de conséquences intéressantes : le noyau de $\mathscr{L}_{|\mathscr{O}_K^*}$ serait alors fini (prendre $B = \{0\}$), et l'image de ce même morphisme serait discrète dans $\mathbb{R}^{r_1+r_2}$, par un procédé bien connu des mordus de topologie. On en déduirait alors que $\mathscr{L}(\mathscr{O}_K^*) \simeq \mathbb{Z}^s$ avec $s \leqslant r_1 + r_2$, et même $s \leqslant r_1 + r_2 - 1$ parce que le plongement logarithmique est inclus presque trivialement dans un hyperplan H (grâce à (2) et à $\ln(|\mathrm{N}(\alpha)|) = 0$) de $\mathbb{R}^{r_1+r_2}$, celui défini par $\sum_i x_i = 0$. Enfin, un résultat classique de théorie des groupes (démontré par exemple dans [**Rau**]) donnerait, comme l'image est libre,

$$\mathscr{O}_K^* \simeq \ker(\mathscr{L}_{|\mathscr{O}_K^*}) \times \mathrm{im}(\mathscr{L}_{|\mathscr{O}_K^*}) \simeq \ker(\mathscr{L}_{|\mathscr{O}_K^*}) \times \mathbb{Z}^s,$$

et il n'y a plus qu'à se convaincre que c'est à peu près ce qu'on veut : le noyau du morphisme est bel et bien exactement l'ensemble des racines de l'unité de K. En effet, si un élément est dans le noyau qui est d'ordre fini, alors il est d'ordre fini et vérifie $x^n = 1$ pour un certain n. L'inclusion réciproque est immédiate.

Pour montrer que $\{x \in \mathscr{O}_K^* \quad | \quad \mathscr{L}(x) \in B\}$ est fini pour tout compact B, on remarque que sous telle condition, il existe $c > 1$ tel que tout x de cet ensemble vérifie $c^{-1} \leqslant |\sigma_i(x)| \leqslant c$. Ainsi, les fonctions symétriques élémentaires des $\sigma_i(x)$ sont bornées et dans $\mathbb{Z}$, donc ne se permettent qu'un nombre fini de valeurs. Alors, il y a un nombre fini de polynômes caractéristiques possibles pour x, donc un nombre fini de valeurs pour x ! □

Bref, le théorème des unités de Dirichlet dit précisément qu'un élément entier inversible est de la forme $ua_1^{n_1} \cdots a_r^{n_r}$ avec u une racine de l'unité, et a_i un générateur libre pour tout i.

Un résultat du même acabit existe dans un cadre plus général :

Théorème 2.24 **(Théorème des unités généralisé)**. — *Si A est un anneau et un $\mathbb{Z}$-module de type fini, alors $(A^*, \cdot)$ est un groupe abélien de type fini.*

Mais il ne s'appliquera pas à $\mathscr{O}_{K,S}$, où S est un ensemble fini d'idéaux premiers de $\mathscr{O}_K$, et

$$\mathscr{O}_{K,S} = \{x \in K | \forall \mathfrak{p} \notin S, \mathrm{ord}_{\mathfrak{p}}(x) \geqslant 0\},$$

qui pourtant vérifie la conclusion du théorème (il est de rang au plus $r + |S|$) : l'argument pour $(\mathscr{O}_{K,S}^*, \cdot)$ est que la suite $1 \to \mathscr{O}_K^* \to \mathscr{O}_{K,S}^* \to \mathbb{Z}^{|S|}$ est exacte. C'est, comme cela n'aura échappé à personne, un anneau localisé en le complémentaire des éléments de S. Par les remarques du début du mémoire, on peut « rendre » $\mathscr{O}_K$ principal par rapport aux idéaux premiers dans la décomposition des générateurs du groupe des classes d'idéaux de $\mathscr{O}_K$.

2.3. En géométrie projective. — À présent, je présente le terrain de jeu des courbes elliptiques, ce qui nécessite quelques prérequis de géométrie projective. Je note $\mathbb{A}^n$ l'espace affine K^n. Mais surtout, puisque je parle de géométrie projective, je dois parler de l'espace projectif, qui est l'objet de la

Définition 2.25 **(Espace projectif)**. — On appelle espace projectif de dimension n, noté $\mathbb{P}^n(K)$, l'ensemble $K^{n+1} \setminus \{0\}$ muni de la relation d'équivalence $\vec{x} \sim \vec{y} \Leftrightarrow \vec{x}$ et $\vec{y}$ sont colinéaires.

On l'appelle aussi *espace des droites vectorielles* (chaque point de l'espace projectif correspond à une droite).

Dans l'espace affine, la plupart des objets géométriques (courbes, surfaces...) sont définis comme étant le lieu des zéros d'un polynôme (ou d'une famille de polynômes). Mais dans le cas projectif, on peut rencontrer un problème si la nullité dépend du choix du représentant. C'est pourquoi on travaille essentiellement avec des polynômes *homogènes*, c'est-à-dire des polynômes qui vérifient, pour un certain d, $P(\lambda x_1, \ldots, \lambda x_n) = \lambda^d P(x_1, \ldots, x_n)$ pour tous λ, tout $\vec{x} = (x_1, \ldots, x_n)$. Pour de tels polynômes, la nullité d'un élément de l'espace projectif est bien définie, puisqu'elle ne dépend plus du représentant.

Se représenter $\mathbb{P}^n(K)$ n'est pas si difficile que ça, puisqu'on peut montrer facilement que $\mathbb{P}^n(K)$ ressemble à $\mathbb{A}^n \sqcup \mathbb{P}^{n-1}(K)$. En fait :

Proposition 2.26. — *$\mathbb{P}^n(K)$ peut être vu comme $\mathbb{A}^n \sqcup \mathbb{A}^{n-1} \sqcup \cdots \sqcup \mathbb{A}^0$.*

On peut détailler cette observation dans le cas du plan projectif $\mathbb{P}^2(K)$: la bijection est l'application

$$f : \begin{cases} \mathbb{P}^2(K) & \to & \mathbb{A}^2 \sqcup \mathbb{P}^1(K) \\ [a,b,c] & \mapsto & \begin{cases} (a/c, b/c) \in \mathbb{A}^2 \text{ si } c \neq 0, \\ [a,b] \in \mathbb{P}^1(K) \text{ sinon.} \end{cases} \end{cases}$$

Et dans ce cas, $\mathbb{P}^1$ représente tous les points à l'infini, un point pour chaque direction. Ainsi, si D est une droite projective d'équation $\alpha X + \beta Y + \gamma Z = 0$, avec α et β non simultanément nuls, alors $[a, b, c] \in D$ avec $c \neq 0$ est envoyé sur le point $(a/c, b/c) \in D'$ d'équation $\alpha x + \beta y + \gamma = 0$, et $[-\beta, \alpha, 0] \in D$ est envoyé sur le point $[-\beta, \alpha] \in \mathbb{P}^1(K)$, qui correspond à la direction de la droite D'. Ceci marche pour toutes les droites, sauf celle d'équation $Z = 0$, qui est envoyée sur la droite contenant tous les points à l'infini de $\mathbb{A}^2 \sqcup \mathbb{P}^1$. Plus généralement, si une courbe projective est définie par $F(X, Y, Z) = 0$, ses points $[a, b, c]$ où $c \neq 0$ sont envoyés sur la courbe affine $f(x, y) = F(x, y, 1) = 0$, et ses points $[a, b, 0]$ sont envoyés sur les points de la droite projective $[a, b]$, et ils correspondent à ce qui pourrait intuitivement être des points qui « manquent » à $f(x, y) = 0$, car à l'infini (ceci peut se vérifier sur des exemples, pour une hyperbole par exemple ces points sont donnés par les directions des asymptotes).

$\mathbb{P}^2(K)$ est donc une sorte de « complétion » de l'espace affine $\mathbb{A}^2$, puisqu'on ajoute dans chaque direction un point à l'infini qui permet d'avoir un résultat géométrique comme dans les rêves : deux droites quelconques se coupent toujours en un et un seul point. Et un fait fantastique réside dans le théorème suivant, qui montre qu'en permettant aux droites d'avoir exactement un point d'intersection, un résultat analogue se généralise à tous les types de courbes :

***Théorème 2.27* (Théorème de Bézout).** — *Soient $\mathscr{C}_1$ et $\mathscr{C}_2$ deux courbes du plan projectif sur un corps algébriquement clos, sans composante commune, de degrés respectifs d_1 et d_2. Alors,*

$$\sum_{P \in \mathscr{C}_1 \cap \mathscr{C}_2} \mathrm{mult}(\mathscr{C}_1 \cap \mathscr{C}_2, P) = d_1 d_2.$$

Le sens rigoureux de la « multiplicité » $\mathrm{mult}(\mathscr{C}_1 \cap \mathscr{C}_2, P)$ sera explicité dans ma démonstration. En vérité, dans le cadre de mon mémoire, seuls deux cas particuliers (intersection d'une droite ou d'une conique avec une courbe) interviennent, mais je me suis efforcé de démontrer ce résultat peu simple par amour du sport. Une ébauche de démonstration est en annexe, le reste découlant de raisonnements algébriques relativement basiques [*].

Voici une manière d'utiliser le théorème de Bézout, pour dénombrer des (ou du moins, calculer des dimensions d'espaces de) courbes vérifiant certaines conditions : j'ajoute assez de points pour « déborder » de l'hypothèse du théorème de Bézout, ce qui permettra de considérer une composante commune aux courbes considérées, et jouera sur la dimension de ce qui nous intéresse. Mais avant, remarquons que si N dénote la dimension de l'espace vectoriel des polynômes homogènes de degré d à $n+1$ indéterminées, alors la dimension de l'espace vectoriel de ces mêmes polynômes s'annulant en r points distincts est supérieure ou égal à $N - r$: chaque point impose une combinaison linéaire entre les coefficients. Bref, l'application :

∗. On a juste besoin de savoir que le nombre N de polynômes homogènes de degré d à n indéterminées est $\binom{n}{d}$. Une façon de dénombrer ceci consiste à vérifier que les $\binom{n+d}{d}$ et N vérifient la même relation de récurrence, $u_{n,d} = u_{n,d-1} + u_{n-1,d}$, et les mêmes valeurs initiales.

Proposition 2.28. — *Par cinq points $P_1,\dots, P_5$ du plan projectif, il passe toujours une conique. Si en plus quatre de ces points ne sont jamais alignés, cette conique est unique (à un coefficient multiplicatif près évidemment).*

La stratégie qui va suivre est, je le répète, un classique, que je réutiliserai plus tard.

Démonstration. — Déjà, la dimension des coniques passant en ces cinq points vaut au moins $\binom{4}{2} - 5 = 1$. À présent, je suppose que je peux trouver trois points alignés sur une droite d'équation $D = 0$. Alors, une conique s'annulant en ces cinq points contient D, sinon les deux courbes auraient trois points en commun, ce qui contredirait le théorème de Bézout qui en autorise deux ici, s'il n'y a pas de facteur commun. Alors, la conique a une équation du type $DD' = 0$, avec D' qui passe par les deux points restants, donc est définie uniquement (à coefficient multiplicatif près) ! Si je ne peux pas trouver deux points alignés, je ne me dégonfle pas et je considère un autre point de la droite $D = (P_4, P_5)$ (par exemple), que j'appelle P_6. Alors, *reductio ad absurdum* : si la dimension cherchée est supérieure ou égale à 2, l'ensemble des coniques passant en ces six points est de dimension au moins 1, et il existe donc une conique non triviale Q passant par P_1, …, P_6. Mais alors, toujours par le théorème de Bézout, Q doit contenir D, et s'écrire $Q = DD' = 0$ où D' passerait par $P_1, \dots, P_3$, donc P_1, P_2 et P_3 sont alignés... Ce qui est supposé impossible ! □

Joli, n'est-ce pas ? De quoi créer un gain de popularité de la géométrie chez les étudiants...

3. Introduction aux courbes elliptiques

On y arrive, enfin ! Commençons par définir l'objet principal de notre étude :

Définition 3.1 **(Courbe elliptique)**. — Une courbe elliptique E sur K est une cubique lisse dans le plan projectif de K.

C'est donc une courbe de degré 3. On s'intéresse au degré 3 entre autres parce que les équations diophantiennes [†] de degré 1 (droites) et 2 (coniques) sont déjà connues et souvent faciles à résoudre. On note souvent $E(K)$ les points de E qui sont à coordonnées toutes dans K.

3.1. Loi de groupe. — Bref, on veut des points rationnels. Et il est intéressant de remarquer, avec le point de vue géométrique, qu'il est facile d'obtenir des points rationnels (ou dans K), une fois un point rationnel (ou dans K) repéré : il suffit de considérer l'intersection entre la tangente en ce point et la courbe elliptique étudiée. De même, si on considère l'intersection de la droite passant par deux points rationnels (ou dans K) et de la courbe elliptique. Mais je radote, il est temps pour moi de démontrer l'existence d'une loi de groupe.

Pour commencer, on se fixe un point O quelconque qui sera l'élément neutre 0_E. Comme sur les figures qui vont suivre, pour P et Q deux points de la courbe, on trace

†. Car, je le rappelle, on fait de l'arithmétique, même si ça ne paraît pas évident ici !

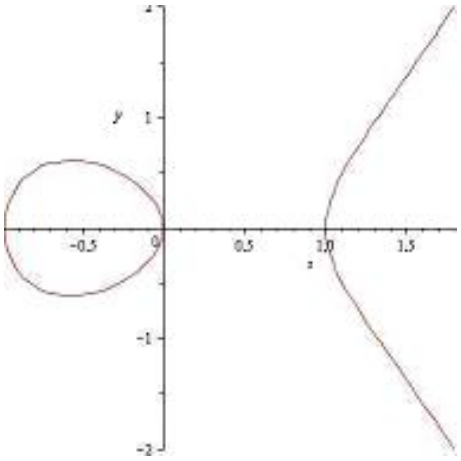

FIGURE 1. Je suis une courbe elliptique, votre humble serviteur.

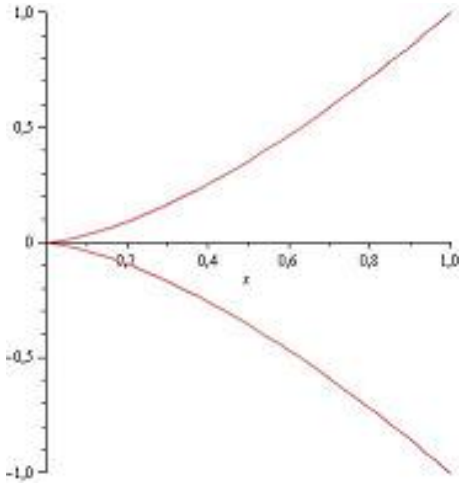

FIGURE 2. Pas moi.

la droite passant par P et Q. En vertu du théorème de Bézout, la droite et la cubique se coupent en un troisième point, noté $P \circ Q$ (on a donc que P, Q et $P \circ Q$ sont alignés). Si $P = Q$, on prend la tangente en P à la place. Ceci ne suffit pas pour avoir notre loi de groupe (à cause de l'élément neutre par exemple), on réitère le procédé avec le point obtenu et O pour avoir $P + Q$.

***Définition 3.2* (Loi de groupe).** — On définit $P + Q = O \circ (P \circ Q)$ et $-P = (O \circ O) \circ P$. On dit alors que l'addition est définie par le procédé de *tangentes et cordes*.

Bref, Ça ressemble, *grosso modo*, à ceci :

Théorème 3.3. — *$(E(\bar{K}), +)$ est un groupe abélien.*

Démonstration. — La stabilité de la somme et de l'inverse se vérifie par construction. $P + 0_E = 0_E \circ (0_E \circ P) = P$, car P, 0_E et $0_E \circ P$ étant alignés et sur E, la droite passant par 0_E et $(0_E \circ P)$ coupe en P. De plus, $P + Q = Q + P$ car on a évidemment $P \circ Q = Q \circ P$, et $P + (-P) = 0_E \circ (P \circ ((0_E \circ 0_E) \circ P))$... On prend son inspiration, on calcule calmement (en se fiant à l'intuition géométrique), pour trouver

$$P + (-P) = 0_E \circ (0_E \circ 0_E) = 0_E.$$

Facile ! Le plus dur à vérifier est l'associativité, qui se vérifie pourtant empiriquement :

Pour l'associativité, on a besoin d'un lemme visiblement technique :

Lemme 1. — *Soient $P_1, \ldots, P_8 \in \mathbb{P}^2(\bar{K})$ distincts. Supposons que 4 d'entre eux ne sont jamais alignés et que 7 d'entre eux n'appartiennent jamais à une même conique. Alors l'espace vectoriel des polynômes homogènes de degré 3 s'annulant en $P_1, \ldots, P_8$ est de dimension 2.*

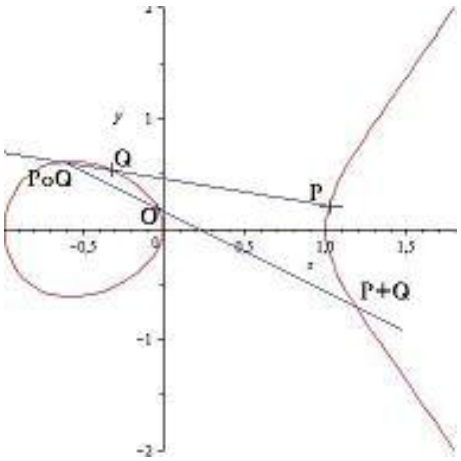

FIGURE 3. Loi de groupe sur E.

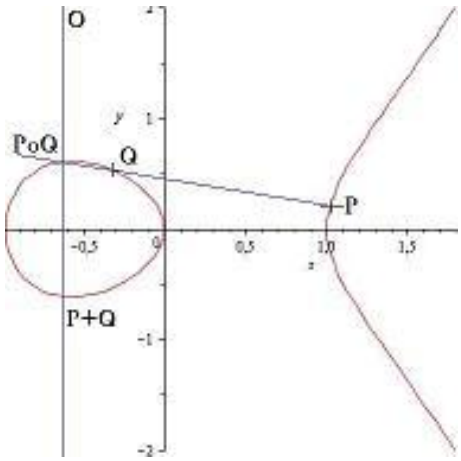

FIGURE 4. O est à l'infini ? Pas de panique !

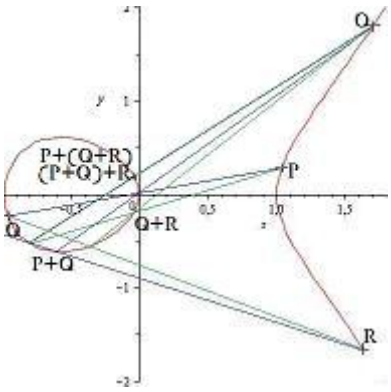

FIGURE 5. Le plus dur ? Ça ne se voit pas...

L'hypothèse n'est pas anodine : sous l'hypothèse du théorème de Bézout, l'intersection d'une cubique et d'une droite contient trois points exactement, et l'intersection d'une cubique et d'une conique contient six points exactement. On peut sentir le lien !

Démonstration du lemme. — C'est un *classic-Bézout*, comme dans la proposition 2.28. Plus concrètement : je sais déjà, par les prérequis de géométrie projective, que la dimension cherchée n est supérieure ou égale à $10 - 8 = 2$. À présent, si (au mieux) je trouve trois points alignés, P_1, P_2 et P_3 disons, soit P_9 sur la même droite d'équation $D = 0$. Toute cubique qui s'annule sur les neuf points en présence contient donc D, et a pour équation $DQ = 0$ où Q est une conique qui s'annule sur $P_4, \ldots, P_8$. Or, par cinq points dont quatre ne sont pas alignés, il ne passe qu'une seule conique Q_0, donc

LQ est un multiple que LQ_0. Bref, la dimension de l'espace des cubiques s'annulant en ces neuf points est 1, donc $n \leqslant 1 + 1 = 2$ et $n = 2$. Si, maintenant, je suppose qu'on n'a jamais trois points alignés, mais qu'en revanche je peux trouver six points (disons P_1,..., P_6) sur une même conique Q, je procède de même en considérant P_9 sur cette conique. Alors, les cubiques s'annulant sur ces neuf points sont encore de la forme $DQ = 0$, pour les mêmes raisons, où D est la droite passant par P_7 et P_8. Ceci constitue encore un espace vectoriel de dimension 1, donc $n \geqslant 2$ et $n = 2$. Si, enfin, il est impossible de trouver trois points alignés ou six points « coconiques », il ne faut pas avoir peur d'en faire trop : considérons deux points P_9, P_{10} sur la droite passant par P_1 et P_2, d'équation $D = 0$. Alors $n = 2$, car si $n \geqslant 3$, alors l'espace des cubiques s'annulant sur ces dix points est de dimension au moins $3 - 2 = 1 > 0$. C'est impossible, car alors ces cubiques s'écriraient $DQ = 0$, et Q serait une conique passant par six points : exclu ! □

On n'est pas encore tirés d'affaire... Ce lemme sert à prouver le vrai lemme qui nous intéresse :

***Lemme 2* (Théorème des neuf points).** — *Soient P_1, ..., P_9 les neuf points communs à deux cubiques $\mathscr{C}_1$ et $\mathscr{C}_2$ dont l'une au moins est irréductible. Si les huit premiers points sont distincts, et qu'une cubique $\mathscr{C}$ passe par ces huit points, alors elle passe par le neuvième.*

Démonstration du lemme : Si par exemple $\mathscr{C}_1$ est irréductible, alors les points qu'elle contient sont ni alignés, ni « coconiques » (à cause du théorème de Bézout, *classic*). D'après le lemme précédent, l'espace vectoriel des cubiques s'annulant sur $P_1, \ldots, P_8$ est de dimension 2, et est donc engendré par les équations de $\mathscr{C}_1$ et $\mathscr{C}_2$. □

L'associativité résultera de ce théorème finement employé. Vous vous en doutez, les neuf points seront les divers points créés par le procédé de tangentes et cordes... Faisons le compte : pour $(P+Q)+R$, les points en jeu sont $P \circ Q$, $O \circ (P \circ Q) = O_1$, $O_1 \circ R$, $O \circ (O_1 \circ R) = U_1$. Pour $P + (Q + R)$, on a $Q \circ R$, $O \circ (Q \circ R) = O_2$, $O_2 \circ P$, $O \circ (O_2 \circ P) = U_2$.

Démonstration. — Soit $\mathscr{C}_1$ la cubique formée de la réunion des droites (P, Q), $(Q \circ R, O)$ et (R, O_1), et $\mathscr{C}_2$ celle définie de même, mais avec les droites (Q, R), $(P \circ Q, O)$ et (P, O_2). Alors :

$$E \cap \mathscr{C}_1 = \{P, Q, R, O, P \circ Q, O_1, Q \circ R, O_2, O_1 \circ R\}$$

$$E \cap \mathscr{C}_2 = \{P, Q, R, O, P \circ Q, O_1, Q \circ R, O_2, O_2 \circ P\}$$

Le théorème des neuf points assure alors, comme E est irréductible, que dans le cas où les huit premiers points sont distincts, alors $O_1 \circ R = O_2 \circ P$ et $(P+Q)+R = P+(Q+R)$ s'ensuit. Si les points en question ne sont pas distincts, on conclut par densité [*]. □

∗. Dans le cas d'un corps de caractéristique non nulle, on peut avoir à invoquer la topologie de Zariski, qui définit les fermés comme étant le lieu d'annulation de polynômes.

J'ai démontré le résultat avec $E(\bar{K})$ car le théorème de Bézout a été démontré dans un corps algébriquement clos [†]. Mais heureusement, $E(K)$ est également un groupe abélien : si P et Q sont à coordonnées dans K, alors $P \circ Q$ aussi, ceci étant dû essentiellement au fait que si un polynôme de degré 3 a deux racines dans K, la troisième est également dans K (penser aux fonctions symétriques élémentaires !).

Le modèle de Weierstrass. — Reportons notre intérêt sur une cubique particulière : une cubique de Weierstrass. Elle a une expression de la forme

$$Y^2Z = X^3 + aXZ^2 + bZ^3,$$

auquel on associe $\Delta = 4a^3 + 27b^2$. Elle mérite notre intention parce qu'elle est assez simple (a et b suffisent, ce qui ne serait pas le cas pour toute expression polynomiale de degré 3), et parce qu'hormis en caractéristique 2 ou 3, toute cubique possédant un point rationnel peut subir un changement de variable l'écrivant sous la forme ci-dessus. En passant à l'expression affine, on se retrouve avec, comme expression d'une courbe elliptique :

$$y^2 = x^3 + ax + b,$$

la condition de lissité étant vérifiée si et seulement si $\Delta \neq 0$. En effet, si l'on définit $F(x, y) = y^2 - x^3 - ax - b$, alors $\frac{\partial F}{\partial x}(x, y) = \frac{\partial F}{\partial x}(x, y) = 0$ revient à dire que $2y = 0$ et $3x^2 + a = 0$, ou encore que x est racine double du polynôme $x^3 + ax + b$... Cas de figure possible uniquement si $\Delta = 0$, par exemple parce que Δ égale le produit des carrés des différentes différences possibles entre racines.

En fait, le signe de Δ donnant le nombre de racines réelles du polynôme, il permet d'exhiber deux types différents (topologiquement parlant) de courbes elliptiques (figures 6 et 7).

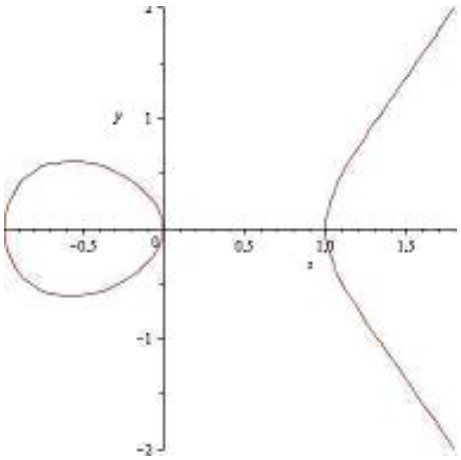

FIGURE 6. $\Delta < 0$, deux composantes connexes.

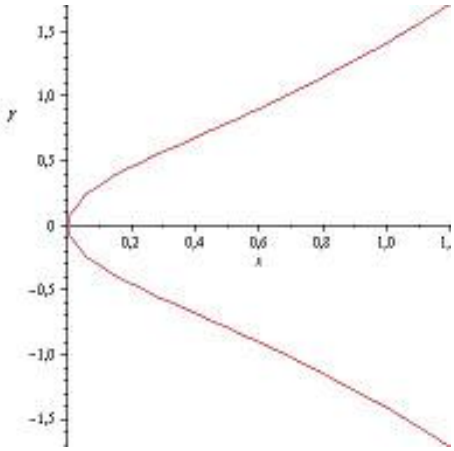

FIGURE 7. $\Delta > 0$, une composante connexe.

†. Un corps a toujours une clôture algébrique. L'idée essentielle consiste en l'emploi du lemme de Zorn sur un ensemble ordonné d'extensions algébriques. Voir [**Jac**].

Dorénavant, on prend pour élément neutre $O = [0, 1, 0]$, le « point à l'infini [‡] » qui est aussi point d'inflexion. Il simplifie donc quelques calculs, et comme la tangente en O coupe la courbe uniquement en O, on a en fait $P + Q = 0_E$ dès que la droite passant par P et Q coupe la courbe en O, c'est-à-dire si cette même droite est, d'après l'interprétation géométrique donnée précédemment, de même direction que l'axe des ordonnées dans le plan affine. Plus précisément, l'opposé de $P = (x, y)$ est $Q = (x, -y)$.

En particulier, si j'appelle α_i les racines de $X^3 + aX + b$, alors $P_i = (\alpha_i, 0)$ est un point de 2-torsion, c'est-à-dire $[2]P_i = 0_E$. Toujours par ce choix de 0_E, on déduit la jolie proposition géométrique suivante : $P + Q + R = 0_E$ si et seulement si P, Q et R sont alignés.

Plus concrètement, on sait calculer explicitement les coordonnées $P + Q$, $-P$, *etc.*, une fois données celles de P et Q (dans le cas d'une cubique de Weierstrass du moins). Il s'agit de résolutions d'équations, où on considère les équations des droites, des tangentes et de la courbe elliptique. Elles ne sont pas très jolies, voire indigestes ; comme on en aura besoin pour la suite, je fournis tout de même quelques expressions de référence :

$$x([2]P) = \frac{x(P)^4 - 2a \cdot x(P)^2 - 8 \cdot x(P) + a^2}{4(x(P)^3 + a \cdot x(P) + b)} \tag{3}$$

$$x(P + Q) + x(P - Q) = \frac{2(x(P) + x(Q))(a + x(P)x(Q)) + 4b}{(x(P) - x(Q))^2} \tag{4}$$

$$x(P + Q)x(P - Q) = \frac{(x(P)x(Q) - a)^2 - 4b(x(P) + x(Q))}{(x(P) - x(Q))^2} \tag{5}$$

où $[2]P = P + P$ (et $[n]P = P + \cdots + P$, n fois).

3.2. Hauteurs. — Pour mon étude des courbes elliptiques, et en particulier le théorème de Mordell-Weil, j'introduis à présent une notion utile qui traduit la « complexité arithmétique » d'un nombre. Pour avoir une idée de la marche à suivre, comparons les nombres $\frac{1}{7}$ et $\frac{142857}{1000000}$. Ils sont assez proches, mais le second est plus compliqué en un certain sens. Pour $x = \frac{a}{b} \in \mathbb{Q}$ (avec a et b premiers entre eux), il est donc naturel de mesurer la complexité à l'aide d'une fonction (dite de hauteur) $H(x) = \max(|a|, |b|)$, et de même on peut définir, dans $\mathbb{P}^n(\mathbb{Q})$, une hauteur via $H(P) = \max(|a_0|, \ldots, |a_n|)$, où $a_0, \ldots, a_n$ sont les coordonnées de P choisies de sorte qu'elles soient premières entre elles. Le problème est que cette définition n'est pas prolongeable à des nombres d'une extension finie K de $\mathbb{Q}$. Je montre donc comment définir une hauteur sur K, pour ensuite prouver que ce n'est pas un objet d'étude anecdotique dans ce cadre.

‡. Je dis *le*, car c'est en effet le seul point projectif de la courbe de Weierstrass dont la troisième coordonnée est nulle.

Valeurs absolues, ou places. — Je propose de trouver dans les lignes à venir une expression de la hauteur qui soit « intrinsèque », au sens où elle ne dépend pas de la forme particulière des éléments de $\mathbb{Q}$, et peut se généraliser à des extensions K sans trop de difficulté. Tout d'abord, je rappelle ce qu'est une valeur absolue sur un corps K :

***Définition 3.4* (Valeur absolue).** — Une valeur absolue v sur un corps K est une application $x \mapsto |x|_v$ de K dans $\mathbb{R}_+$ telle que pour tous $x, y \in K$,

1. $|x|_v = 0$ si et seulement si $x = 0$,
2. $|xy|_v = |x|_v |y|_v$,
3. il existe $C_v > 0$ (indépendant de x et y) tel que $|x + y|_v \leqslant C_v \max(|x|_v, |y|_v)$.

Dans le cas où on peut prendre $C_v = 1$, la valeur absolue est dite *ultramétrique*. Dans le cas $C_v = 2$, la valeur absolue vérifie l'inégalité triangulaire (ce n'est pas trivial !), et réciproquement. Les valeurs absolues non ultramétriques sont dites archimédiennes dans mon cadre d'étude.

Si on remplace $|\cdot|$ par $|\cdot|^\alpha$ avec $\alpha > 0$, on obtient encore une fois une valeur absolue (même si l'inégalité triangulaire ne se transmet pas très bien, le cas échéant). Si deux valeurs absolues sont reliées par une puissance, on dit qu'elles sont équivalentes, et les classes d'équivalence sont appelées les *places*, et forment l'ensemble M_K. Les places archimédiennes forment $M_{K,\infty}$.

Dans le cas de $\mathbb{Q}$, les places sont représentées par les valeurs absolues p-adiques (pour p premier) et la valeur absolue classique. Ces valeurs absolues p-adiques sont définies, si $x = p^n \frac{a}{b}$ avec a, b et p premiers entre eux deux à deux, par $|x|_p = p^{-n}$ (sauf si $x = 0$, auquel cas on pose $|0|_p = 0$). Le fait qu'il n'y en ait pas d'autres découle du théorème d'Ostrowski, auquel on reviendra plus tard ; dans ce lot de valeurs absolues, la valeur absolue standard (ici notée $|\cdot|_\infty$) paraît un peu « bizarre ». En construisant les places sur un corps de nombres K, j'expliquerai cette étrangeté, en plus de définir les hauteurs sur un corps K puis sur une courbe elliptique.

***Théorème 3.5* (Formule du produit sur $\mathbb{Q}$).** — *Si $x \in \mathbb{Q}^*$, alors*

$$\prod_{v \in M_\mathbb{Q}} |x|_v = 1.$$

On déballe les définitions des différents $|x|_v$ et la démonstration s'écrit toute seule. L'analogie pour K sera plus intéressante, mais la formule du produit sur $\mathbb{Q}$ permet déjà d'établir ceci :

Corollaire 3.6. — *Pour $P = [x_0, \ldots, x_n] \in \mathbb{P}^n(\mathbb{Q})$, on a*

$$H(P) = \prod_{v \in M_\mathbb{Q}} \max(|x_0|_v, \ldots, |x_n|_v).$$

D'après la formule du produit, si $\vec{y} = \lambda \vec{x}$, alors $\prod_{v \in M_\mathbb{Q}} |\lambda|_v = 1$, donc le membre de droite est le même pour $\vec{x}$ et $\vec{y}$, et l'égalité est sensée.

Démonstration. — quitte à changer les coordonnées projectives, on peut supposer les x_i premiers entre eux. Alors, $\max(|x_0|_p, \dots, |x_n|_p) = 1$ pour tout premier p, et le membre de droite égale $\max(|x_0|_\infty, \dots, |x_n|_\infty) = H(P)$. □

Alors, une piste pour étendre H à un corps de nombres consiste à explorer les places de K. Soient r_1 et r_2 les entiers définis en géométrie des nombres (avec $r_1 + 2r_2 = n$). Alors,

Définition 3.7. — Pour chaque plongement σ, je note $|x|_\sigma = |\sigma(x)|$ si σ est réel, $|x|_\sigma = |\sigma(x)|^2$ sinon.

Ça nous fait $r_1 + r_2$ places, qui forment $M_{K,\infty}$ en fait. En s'inspirant du cas de $\mathbb{Q}$, on peut également songer à :

Définition 3.8. — Pour chaque $\mathfrak{p}$ idéal premier de $\mathscr{O}_K$, je note $|x|_\mathfrak{p} = \mathrm{N}(\mathfrak{p})^{-\mathrm{ord}_\mathfrak{p}(x)}$.

Si $\mathfrak{p}$ est un idéal premier de $\mathscr{O}_K$, alors $\mathfrak{p} \cap \mathbb{Z}$ est un idéal premier de $\mathbb{Z}$ (parce que $\mathbb{Z}/(\mathfrak{p} \cap \mathbb{Z})$ s'injecte dans $\mathscr{O}_K/\mathfrak{p}$ qui est intègre). Je peux donc désigner p le nombre premier tel que $\mathfrak{p} \cap \mathbb{Z} = p\mathbb{Z}$. Inversement, $p\mathscr{O}_K$ n'est pas forcément un idéal premier, et se factorise en $\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_g^{e_g}$. Alors, $\mathrm{N}(p) = p^n$, donc $\mathrm{N}(\mathfrak{p}_i) = p^{f_i}$ pour un certain f_i (en outre, $\sum_i e_i f_i = n$). Dans les deux cas, on dit alors que $|\cdot|_\mathfrak{p}$ est une place de K au-dessus de p, ce qui est noté $\mathfrak{p}|p$.

Par ce procédé, on obtient toutes les places ! En effet :

Théorème 3.9 **(Théorème d'Ostrowski)**. — *Sur un corps de nombres K, les seules places sont celles des valeurs absolues associées aux idéaux premiers de $\mathscr{O}_K$ et aux plongements de K dans $\mathbb{C}$.*

Ce théorème démontré dans [**Bou**] n'est pas très utile, mais permet d'avoir la satisfaction de savoir qu'on n'oublie pas de places par notre procédé, et d'espérer que la formule du produit reste valable dans K. Dans ce cadre là, la valeur absolue classique sur $\mathbb{Q}$ est simplement celle associée au plongement $\begin{array}{ccc} \mathbb{Q} & \to & \mathbb{C} \\ x & \mapsto & x \end{array}$.

Il découle de tout ceci que, pour $x \in K$,

$$\prod_{\mathfrak{p}|p} |x|_\mathfrak{p} = |\mathrm{N}(x)|_p \text{ et } \prod_{\sigma : K \hookrightarrow \mathbb{C}} |x|_\sigma = |\mathrm{N}(x)|_\infty.$$

En effet,

$$N(x) = \pm N(x\mathscr{O}_K) = \pm \prod_\mathfrak{p} N(\mathfrak{p})^{-\mathrm{ord}_\mathfrak{p}(x)} = \pm \prod_p p^{\sum_{\mathfrak{p}|p} f_\mathfrak{p} \mathrm{ord}_\mathfrak{p}(x)}.$$

Alors,

$$|\mathrm{N}(x)|_p = p^{-\sum_{\mathfrak{p}|p} f_\mathfrak{p} \mathrm{ord}_\mathfrak{p}(x)} = \prod_{\mathfrak{p}|p} p^{-f_\mathfrak{p} \mathrm{ord}_\mathfrak{p}(x)} = \prod_{\mathfrak{p}|p} \mathrm{N}(\mathfrak{p})^{-\mathrm{ord}_\mathfrak{p}(x)} = \prod_{\mathfrak{p}|p} |x|_\mathfrak{p}.$$

Pour $|\mathrm{N}(x)|_\infty$, c'est immédiat d'après (2).

Alors, $\prod_{w\in M_K} |x|_w = \prod_{v\in M_\mathbb{Q}} \prod_{w|v} |x|_w$, où $w|v$ signifie ce qui a été dit pour $\mathfrak{p}|p$ dans le cas d'un idéal premier, et les plongements sont mis ensemble au-dessus de la valeur absolue classique. Ce calcul se simplifie pour donner $\prod_{v\in M_\mathbb{Q}} |\mathrm{N}(x)|_v = 1$. Bref, je viens d'établir :

***Théorème 3.10* (Formule du produit sur K).** — *Si $x \in K^*$, alors*

$$\prod_{v\in M_K} |x|_v = 1.$$

D'où :

***Définition 3.11* (Hauteur sur K).** — Pour $P = [x_0, \dots, x_n] \in \mathbb{P}^n(K)$, on définit

$$H_K(P) = \prod_{v\in M_K} \max(|x_0|_v, \dots, |x_n|_v).$$

Au vu de la définition, on peut se demander ce qu'il se passe quand on passe de H_K à H_L avec $K \subseteq L$. On obtient simplement $H_L(P) = H_K(P)^{[L:K]}$ par des manipulations proches de celles qui précèdent, en raisonnant sur les places de L au dessus de K, et en se rappelant que $\sum_v e_v f_v = [L : K]$. Ceci motive la définition suivante :

***Définition 3.12* (Hauteur absolue).** — $H : \mathbb{P}^n(\bar{\mathbb{Q}}) \to \mathbb{R}$ est définie par

$$H(P) = H_K(P)^{1/[K:\mathbb{Q}]}$$

si $P \in \mathbb{P}^n(K)$.

Pour $\alpha \in K$, la hauteur de α est définie comme la hauteur de $[1, \alpha] \in \mathbb{P}^1(K)$. Comme on peut s'y attendre, il est possible de lier la hauteur au polynôme minimal d'un nombre algébrique. En effet :

***Proposition 3.13*.** — *Soit α un nombre algébrique, $K = \mathbb{Q}(\alpha)$. Si son polynôme minimal s'écrit dans $\mathbb{Z}[X]$ sous la forme $P = a\prod_{i=1}^{d}(X - \alpha_i)$, alors :*

$$H_K(\alpha) = |a| \prod_{i=1}^{d} \max(1, |\alpha_i|).$$

Pour relier H_K aux coefficients du polynôme minimal, j'aurai besoin de :

***Lemme 3* (Lemme de Gauss).** — *Soient $P, Q \in K[X]$, et $\|P\|_v$ la norme supérieure des coefficients de P, pour v valeur absolue ultramétrique. Alors,*

$$\|PQ\|_v = \|P\|_v \|Q\|_v$$

On parle de lemme de Gauss, parce qu'il généralise le cas $\mathbb{Q}$ (les normes p-adiques expriment mieux la situation que le contenu, qui a le défaut de nécessiter des coefficients entiers).

Démonstration du lemme de Gauss. — Si $\|P\|_v = \|Q\|_v = 1$, alors $P, Q \in \mathscr{O}_{K,v}[X]$ et, mieux, $P, Q \neq 0 \mod \alpha\mathscr{O}_{K,v}[X]$ où α est un générateur de l'unique [(*)] idéal maximal de $\mathscr{O}_{K,v}$. Comme le quotient est intègre, $PQ \neq 0 \mod \alpha\mathscr{O}_{K,v}[X]$. Donc $\|PQ\|_v = 1$. Si $\|P\|_v$ ou $\|Q\|_v$ n'égalent pas 1, on peut factoriser : $P = \alpha^m P'$, $Q = \alpha^n Q'$, et le tour est joué. □

Démonstration de la proposition. — Soit $L = K(\alpha_1, \dots, \alpha_d)$, déterminons H_L qui est plus simple à calculer : Armé du lemme de Gauss, on peut immédiatement écrire, pour $\mathfrak{q} \in M_L$, que $1 = \|P\|_\mathfrak{q} = |a|_\mathfrak{q} \prod_{i=1}^{d} \max(1, |\alpha_i|_\mathfrak{q})$, la première égalité étant due au fait que si tel n'était pas le cas, on pourrait encore factoriser P qui ne serait plus irréductible. On a alors, en faisant le produit sur tous les $\mathfrak{q} \in M_L$ (je rappelle que je veux obtenir H_L) :

$$1 = \prod_{\mathfrak{q} \in M_L} |a|_\mathfrak{q} \prod_{i=1}^{d} \prod_{\mathfrak{q} \in M_L} \max(1, |\alpha_i|).$$

D'une part, la formule du produit donne :

$$\prod_{\mathfrak{q} \in M_L} |a|_\mathfrak{q} \prod_{\sigma : L \hookrightarrow \mathbb{C}} \underbrace{|a|_\sigma}_{= |\sigma(a)|^{1,2} = |a|^{1,2}} = 1,$$

donc $\prod_{\mathfrak{q} \in M_L} |a|_\mathfrak{q} = |a|^{-[L:\mathbb{Q}]}$. D'autre part, $|\alpha_i|_\mathfrak{q} = |\sigma_i(\alpha)|_\mathfrak{q}$, or $\sigma_i(\alpha)\mathscr{O}_L = \alpha\mathscr{O}_L$, donc $|\alpha_i|_\mathfrak{q} = |\alpha|_\mathfrak{q}$. Bref,

$$H_L(\alpha) = \prod_{\mathfrak{q} \in M_L} \max(1, |\alpha|_\mathfrak{q}) \prod_{\sigma : L \hookrightarrow \mathbb{C}} \max(1, |\alpha|_\sigma) = |a|^{[L:\mathbb{Q}]/d} \left(\prod_{i=1}^{d} \max(1, |\alpha_i|) \right)^{[L:K]},$$

car $\prod_{\sigma : L \hookrightarrow \mathbb{C}} \max(1, |\alpha|_\sigma) = \prod_{\eta : K \hookrightarrow \mathbb{C}} \max(1, |\alpha|_\eta)^{[L:K]} = \left(\prod_{i=1}^{d} \max(1, |\alpha_i|) \right)^{[L:K]}$, la première égalité étant due au lemme de Dedekind sur le nombre de prolongements des morphismes de corps. Alors, en prenant les racines $[L : K]$-ièmes [(†)], on a H_K. □

Cette hauteur trouve une de ses principales forces dans le théorème de finitude suivant :

Théorème 3.14 **(Théorème de finitude de Northcott et Kronecker)**
Soit $d \geqslant 1$, $X > 0$. Alors, $\{P \in \mathbb{P}^n(\bar{\mathbb{Q}}) | [\mathbb{Q}(P) : \mathbb{Q}] \leqslant d, H(P) \leqslant X\}$ est fini. De plus, on a $H(P) = 1$ si et seulement si P possède des coordonnées projectives égales toutes à zéro ou une racine de l'unité.

Seule la première partie, celle due à Northcott [(‡)], m'importe vraiment pour la suite. Je ne démontrerai pas la seconde partie, bien que simple.

∗. $\mathscr{O}_{K,v}$ est un anneau localisé en v, donc est local et a un *unique* idéal maximal.

†. N'oublions pas que $d = [K : \mathbb{Q}]$, donc $[L : \mathbb{Q}]/d = [L : K]$ par multiplicativité des degrés.

‡. J'ai bien révisé quelle partie est de Northcott et quelle partie est de Kronecker ! Mais ne me posez pas la même question pour le principe de Huyguens-Fresnel...

Démonstration. — Cette formulation rend le plus problème plus compliqué, en fait si $P = [1, \alpha_1, \ldots, \alpha_n]$ (quitte à permuter les coordonnées le 1 est au début) avec les α_i dans un corps de nombres K « assez grand », on vérifie trivialement que

$$H(\alpha_i) \leqslant H(P) \text{ et } [\mathbb{Q}(\alpha_i) : \mathbb{Q}] \leqslant [\mathbb{Q}(P) : \mathbb{Q}].$$

Alors, j'ai juste à montrer que $\{\alpha \in \bar{\mathbb{Q}} | [\mathbb{Q}(\alpha] : \mathbb{Q}] \leqslant d, H(\alpha) \leqslant X\}$ est fini ; la condition sur le degré $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ du polynôme minimal et sur la hauteur donnent, d'après la proposition précédente, une borne sur les coefficients (entiers) de ce même polynôme, d'où la finitude. □

Au sujet des hauteurs, il est possible de générer beaucoup d'inégalités à leur sujet, en partant de la proposition qui suit :

Proposition 3.15. — *Soient $P_0, \ldots, P_m$ des polynômes homogènes de degré d en $(x_0, \ldots, x_n)$, et notons Z le lieu des zéros communs des P_i.*
Soit $\phi : \begin{cases} \mathbb{P}^n \smallsetminus Z & \to & \mathbb{P}^m \\ \vec{x} & \mapsto & (P_0(\vec{x}), \ldots, P_m(\vec{x})) \end{cases}$ *. Alors :*

- *Il existe une constante C_1 telle que pour tout $\vec{x} \in (\mathbb{P}^n \smallsetminus Z)(\bar{\mathbb{Q}})$,*
$$H(\phi(\vec{x})) \leqslant C_1 H(\vec{x})^d. \tag{6}$$
- *Si $V \subseteq \mathbb{P}^n$ est une sous-variété fermée (c'est-à-dire est le lieu des zéros de Q_i homogènes) ne rencontrant pas Z, alors il existe C_1 et C_2 tels que pour tout $\vec{x} \in V(\bar{\mathbb{Q}})$,*
$$C_2 H(\vec{x})^d \leqslant H(\phi(\vec{x})) \leqslant C_1 H(\vec{x})^d. \tag{7}$$

Avant de démontrer cette proposition, méditons sur les propos d'un poète inconnu, Sergueï Bernikov :

> De démonstrations, je veux plus qu'un *ersatz* !
> Des polynômes pris de semblables degrés,
> Qui dans un même lieu se voient désintégrés :
> Je sens comme un parfum pur de *Nullstellensatz*...

Démonstration. — Pour considérer la hauteur de $\vec{x}$ et $\phi(\vec{x})$ fixés, on se place dans $K = \mathbb{Q}(x_0, \ldots, x_n)$, et alors $H_K(\vec{x}) = \prod_{v \in M_K} \max(|x_0|_v, \ldots, |x_n|_v)$. Pour alléger les notations, je noterai dans les sommes $\vec{x}^i = {x_0}^{i_0} \cdots {x_n}^{i_n}$ avec $i_0 + \cdots + i_n = i$. Si $P_i(\vec{x})$ s'écrit $\sum_j a_j^{(i)} \vec{x}^j$, alors :

$$|P_i(\vec{x})|_v \leqslant \underbrace{\binom{n+d}{d}_v}_{=N_v} \max_j |a_j^{(i)}|_v \max_i |x_i|_v^d$$

où l'apparition du coefficient binomial est dû au nombre maximal N de monômes possibles à $n+1$ indéterminées, de degré d, comme j'en ai déjà parlé. Gardons également à l'esprit que ces valeurs absolues ne vérifient pas l'inégalité triangulaire classique mais une inégalité dépendant de v, d'où l'annotation v en indice de N. Bref, comme $\max_j |a_j^{(i)}|_v = 1$ sauf pour un nombre fini de places (les plongements sont en nombre

fini, et seul un nombre fini d'idéaux premiers intervient dans la factorisation des $a_j^{(i)}\mathcal{O}_K$), le produit $\prod_v A_v$ sera fini. Alors :

$$H_K(\phi(\vec{x})) = \prod_v \max_i |P_i(\vec{x})| \leqslant \prod_v N_v A_v \max_i |x_i|_v^d = \left(\prod_v N_v A_v\right) H_K(\vec{x})^d,$$

d'où la première inégalité en prenant pour C_1 la racine $[K:\mathbb{Q}]$-ième de $\left(\prod_v N_v A_v\right)$.

Pour montrer la seconde inégalité, la bonne idée nous est soufflée : Comme le polynôme X_j s'annule, pour tout j, sur le lieu commun des zéros des Q_i et des P_i (en $\vec{0}$ par exemple), par le *Nullstellensatz* de Hilbert appliqué dans $\bar{\mathbb{Q}}$, il existe donc $M \in \mathbb{N}^*$ et des $A_i^{(j)}$, $B_i^{(j)}$ tels que

$$X_j^M = \sum_{i=0}^{m} A_i^{(j)} P_i + \sum_{i=0}^{r} B_i^{(j)} Q_i. \tag{8}$$

On peut supposer les A_i homogènes de degré $M-d$ et à coefficiens dans $K[X]$ pour K « assez grand », et alors on obtient, pour $\vec{x} \in V$:

$$x_j^M = \sum_{i=0}^{m} A_i^{(j)} P_i(\vec{x}),$$

d'où :

$$|x_j|_v^M \leqslant (m+1)_v \max_i |A_i^{(j)}|_v \max |P_i(\vec{x})|_v \leqslant A_v^{'} |x_i|_v^{M-d} \max_i |P_i(\vec{x})|_v.$$

Comme toujours, $A_v^{'} = 1$ sauf pour un nombre fini de places. Donc l'inégalité $\max_j |x_j|_v^d \leqslant A_v^{'} \max_i |P_i(\vec{x})|_v$ entraine, en faisant le produit sur toutes les places et en prenant la racine $[K:\mathbb{Q}]$-ième, le résultat voulu avec $C_2 = \sqrt[[K:\mathbb{Q}]]{\prod_{v \in M_K} A_v'}$. □

Cet énoncé se résume par l'écriture

$$h(\phi(\vec{x})) = d \cdot h(\vec{x}) + O(1), \tag{9}$$

où h est le logarithme naturel de H (souvent préféré à H), toujours positif ou nul car H est toujours supérieur ou égal à 1. Je vais déduire quelques résultats à venir de cette inégalité très sympathique.

Définition 3.16. — Soit $E(K)$ une courbe elliptique. On définit la hauteur de $P \in E(K)$ par $h(P) = \begin{cases} h(x(P)) & \text{si } P \neq 0_E \\ 0 & \text{sinon.} \end{cases}$

Plus précisément, $h(P) = h([1, x(P)])$ avec $[1, x(P)] \in \mathbb{P}^1(K)$.

Proposition 3.17. — *On a, pour tous $P \in E(K)$, $h([2]P) = 4h(P) + O(1)$.*

Cette proposition pourrait se déduire de celle qui va suivre (à quelques aménagements près pour éviter de se mordre la queue), mais c'est une application simple des inégalités précédentes.

Démonstration. — Dans le cas où $P = 0_E$, il y a rien à dire. De même si $[2]P = 0_E$, mais c'est moins trivial ; dans ce cas, on veut juste $h(P) = O(1)$... Or, le nombre de points de 2-torsion est borné : ils correspondent géométriquement aux points de tangente verticale, et ces points ne foisonnent pas.

Dans le cas général, remarquons que l'inégalité demandée ressemble fortement à la dernière proposition démontrée... Mais pas tout à fait, puisque ϕ doit être un vecteur composé de polynômes homogènes. Inspirons-nous de la *formule de duplication* que je rappelle, qui relie $[2]P$ à P :

$$x([2]P) = \frac{x(P)^4 - 2a \cdot x(P)^2 - 8b \cdot x(P) + a^2}{4(x(P)^3 + a \cdot x(P) + b)}.$$

Face à cette expression, le choix de ϕ se dessine : on prend

$$\phi(T, X) = (4T(X^3 + aXT^2 + bT^3), X^4 - 2aX^2T^2 - 8bXT^3 + a^2T^4),$$

puisqu'alors $\phi(1, x(P)) = [1, x([2]P)]$ (remarquez l'aménagement fait pour avoir des *polynômes* en coordonnées).

Pour vérifier les hypothèses de la proposition 3.15, il faut encore vérifier qu'il n'y a pas de zéro commun entre la courbe elliptique et les polynômes-coordonnées de ϕ. Et un fait curieux qui fait plaisir quand on tombe dessus, c'est que cela marche parce que $\Delta \neq 0$, ce que je n'attendais pas personnellement ! En effet, les divisions euclidiennes successives de $X^4 - 2aX^2 - 8bX + a^2$ par $X^3 + aX + b$ conduisent à un reste de la forme $u\Delta$ avec $u \in K^*$, et ceci induit qu'ils sont premiers entre eux. Mais passons : les calculs donnent

$$h([2]P) = h(1, x([2]P)) = h(\phi(1, x(P))) = 4h(1, x(P)) + O(1) = 4h(P) + O(1).$$

□

Proposition 3.18 **(Presque-loi du parallélogramme)**. — *Pour tous $P, Q \in E(K)$, on a*

$$h(P + Q) + h(P - Q) = 2h(P) + 2h(Q) + O(1),$$

et $h(P) = h(-P)$.

Une fois de plus, un $O(1)$ subsiste et rend le résultat frustrant, malgré le bon usage de la proposition 3.15. J'ai caché la démonstration dans l'annexe, saurez-vous la retrouver ?

Cette hauteur est capable de rendre bien des services, mais en vue du théorème de Mordell-Weil j'aimerais me débarrasser des $O(1)$. À cet effet, je passe à la définition de la hauteur de Néron-Tate, ou hauteur canonique.

Théorème-Définition 1 **(Hauteur de Néron-Tate)**. — *Soit $E(K)$ une courbe elliptique. On définit une hauteur de Néron-Tate par la formule*

$$\hat{h}(P) = \lim_{n \to \infty} \frac{h(x(2^n P))}{4^n}.$$

Alors, $\hat{h}(P) = h(P) + O(1)$, et $\hat{h}$ satisfait la loi du parallélogramme, donc est quadratique. De plus, $\hat{h}(P) = 0$ si et seulement si P est d'ordre fini.

Avec le théorème de Minkowski, on peut même montrer qu'elle est définie positive, mais on n'en a pas besoin.

Ébauche de démonstration. — La convergence de la limite est essentiellement due au fait que la suite est de Cauchy dans $\mathbb{R}$, ceci se déduisant lui-même du fait que $|h([2]) - 4h|$ est majoré. C'est le passage à la limite avec le 4^n qui permet d'avoir une loi du parallélogramme parfaite.

Si $[m]P = 0$ avec $m > 0$, alors $\hat{h}(P) = \frac{1}{m^2}\hat{h}([m]P) = 0$. Réciproquement, si $\hat{h}(P) = 0$, alors pour tout $m \in \mathbb{Z}$, $\hat{h}([m]P) = 0$. Donc $\{[m]P | m \in \mathbb{Z}\}$ est de hauteur bornée, donc est fini, et P est d'ordre fini. □

Tout comme h, la hauteur canonique $\hat{h}$ vérifie la propriété de finitude des ensembles bornés, ce qui me servira tout à l'heure.

3.3. Théorème de Mordell-Weil. — On en arrive à l'énoncé qui a fait l'objet initial de ce mémoire.

***Théorème 3.19* (Théorème de Mordell-Weil).** — *Le groupe $E(K)$ est de type fini.*

En d'autres termes, tous les points rationnels d'une courbe peuvent être obtenus à l'aide de tangentes et de cordes, à partir d'un nombre fini d'entre eux ! Malheureusement (pour moi), la démonstration fut parfois malaisée. En fait, la démonstration que j'ai abordée en premier est une démonstration « naturelle », au sens où on ne peut pas (trop) accuser la démonstration d'user d'astuces et d'être parachutée, d'autant plus qu'elle démontre bien plus que ce que j'annonce dans ce mémoire. Mais elle utilise des outils de théorie de Galois et de cohomologie (il s'agit de cohomologie galoisienne, en fait), dont l'introduction seule aurait pu justifier dix pages de plus dans ce mémoire. J'ai donc opté pour une démonstration « élémentaire », c'est-à-dire qu'elle ne nécessite pas plus de prérequis que ceux de la première partie, et utilise de manière décisive, comme l'autre démonstration, le théorème des unités et la factorisation des idéaux. Le point central de la démonstration est le suivant.

***Théorème 3.20* (Théorème de Mordell-Weil (faible))**
Le groupe $\Gamma = E(K)/2E(K)$ est fini.

Le théorème de Mordell-Weil faible implique le théorème de Mordell-Weil, et par ailleurs des représentants de chaque classe de Γ permettent de déterminer les générateurs de $E(K)$: ils sont parmi les éléments vérifiant $\hat{h}(P) \leqslant C$ où C est la plus haute valeur atteinte par $\hat{h}$ parmi les représentants choisis arbitrairement.

En effet : soit g un élément de $E(K)$. Si $\|g\| \leqslant \sqrt{C}$ (où $\|\cdot\|$ désigne ici la racine carrée de $\hat{h}$), alors g est évidemment dans le sous-groupe engendré par

$$S = \{x \in E(K) \mid \|x\| \leqslant \sqrt{C}\}.$$

Je suppose donc $\|g\| > \sqrt{C}$, et montre qu'il est combinaison linéaire d'éléments de S. Si je pose $g_0 = g$ et $\{y_1, \ldots, y_m\}$ un système de représentants de Γ, alors $g_1 = \frac{g_0 - y_1}{2}$ (écriture possible pour y_1 tel que $g_0 \equiv y_1 \mod 2E(K)$) vérifie $\|g_1\| < \|g_0\|$,

et on utilise décisivement le fait que $\|g_0\| > \sqrt{C}$. Si $\|g_1\| > \sqrt{C}$, on peut encore construire $g_2 = \frac{g_1 - y_2}{2}$ qui vérifie $\|g_2\| < \|g_1\| < \|g_0\|$. Ce procédé ne peut se poursuivre indéfiniment que si on obtient uniquement des $\|g_n\| > \sqrt{C}$, ce qui est impossible puisqu'alors on aurait une infinité d'éléments strictement plus petits que $\|g_0\|$, et la propriété de Northcott assure qu'il y en a un nombre fini ! Ainsi, au bout d'un certain n_0, on a $\|g_{n_0}\| \leqslant \sqrt{C}$, et comme g est combinaison linéaire de g_{n_0} et des y_i qui sont tous dans S, on a le résultat : l'ensemble fini S engendre $E(K)$. Nous voilà rassurés, on peut démontrer le fameux théorème.

Démonstration. — C'est ici que la démonstration ici présentée et la « cohomologique » diffèrent : là où la démonstration « cohomologique » donne un morphisme de groupes $\Gamma \to \mathrm{H}^1(K, E[2])$, ici j'étudie un morphisme de groupes $\Gamma \to (K^*/K^{*2})$. Pourquoi ce choix ? Elle découle de quelques considérations arithmétiques. Supposons que je me place dans un anneau $\mathcal{A}$ suffisamment riche pour pouvoir faire de l'arithmétique (un anneau principal, donc factoriel, serait parfait), dont le corps de fractions serait K. Je pourrais alors écrire x et y sous forme de fractions irréductibles A/B et C/D, avec $A, B, C, D \in \mathcal{A}$. Après avoir reporté ceci dans $y^2 = x^3 + ax + b$ puis enlevé les fractions, puisque je tiens à faire de l'arithmétique comme dans $\mathbb{Z}$ (mais en mieux !), j'obtiendrais :

$$B^3C^2 = D^2 \prod_{i=1}^{3} (A - \alpha_i B),$$

où les α_i sont les racines de $x^3 + ax + b$. Même sans connaissance précise de $\mathcal{A}$, je peux déduire facilement, mais ça viendra plus tard, que

$$C^2 = \prod_{i=1}^{3} (A - \alpha_i E^2),$$

où $E^2 = B$. Si j'arrive à montrer que les facteurs sont premiers entre eux (ce qui viendra aussi plus tard, pour le suspense), alors tous ces éléments sont des carrés, à un élément inversible près. Ou plutôt, ce sont des éléments inversibles, à un carré près. Le groupe $\mathcal{A}^*$ n'a plus qu'à avoir le bon goût d'être de type fini dans le meilleur des cas, puisque qu'alors $\mathcal{A}^*/\mathcal{A}^{*2}$ est fini, et il s'avère qu'on pourra relier Γ à $\mathcal{A}^*/\mathcal{A}^{*2}$ (les deux quotients se ressemblent, après tout) ! Ceux qui ont bien suivi ce mémoire ont peut-être déjà compris quel sera le choix de $\mathcal{A}$ (j'ai mis des pistes), mais je prie de ne pas gâcher la surprise aux autres.

Supposer que les trois racines sont dans K ne pose pas de problème, quitte à remplacer K par $L = K(\alpha_1, \alpha_2, \alpha_3)$. En effet, la véracité du théorème de Mordell-Weil faible pour L induit celle pour K [(*)].

∗. Ce n'est pas évident : il faut remarquer que pour tout P dans le noyau Φ du morphisme naturel $E(K)/2E(K) \to E(L)/2E(L)$, il existe $Q_P \in E(L)$ tel que $[2]Q_P = P$. Alors, on peut construire une injection entre ce noyau et l'ensemble des applications de $\mathrm{Gal}_K(L)$ dans $E[2]$ qui est fini, donc le noyau est fini. La suite exacte $0 \to \Phi \to E(K)/2E(K) \to E(L)/2E(L)$ permet de conclure. Toutefois, si seul le théorème de Mordell nous intéresse ultimement avec K, on peut remarquer que s'il est vrai pour L, alors il est vrai pour K ; il suffit donc de démontrer le théorème de Mordell-Weil faible pour L, sans se soucier de sa justesse pour K.

Ces considérations étant mises de côté pour un instant, définissons les applications suivantes :

$$\psi_i : \begin{cases} E(K) & \to & (K^*/K^{*2})^3 \\ P & \mapsto & \begin{cases} x(P) - \alpha_i & \text{si} \quad P \neq P_i, 0_E \\ (\alpha_i - \alpha_j)(\alpha_i - \alpha_k) & \text{si} \quad P = P_i \\ 1 & \text{si} \quad P = 0_E \end{cases} \end{cases}$$

Et on pose $\psi = (\psi_1, \psi_2, \psi_3)$. Le choix de valeur de ψ en P_i n'est pas tout à fait parachuté : comme

$$(x - \alpha_i) = \frac{y^2}{(x-\alpha_j)^2(x-\alpha_k)^2}(x-\alpha_j)(x-\alpha_k),$$

on a $x - \alpha_i = (x-\alpha_j)(x-\alpha_k) \mod K^{*2}$. Mais d'autres choix auraient été possibles.

On admet (mais on sait montrer) que ψ est un morphisme, c'est-à-dire : pour tout $P, Q \in E(K)$ on a $\psi(P+Q) = \psi(P)\psi(Q)$. Elle découle de considérations géométriques, où on écrit ce qu'il se passe selon les choix de P et Q, et où on réécrit savamment les équations vérifiées par les forces en présence.

Lemme 4. — *On a* $\ker(\psi) = 2E(K)$.

Le fait que $2E(K) \subseteq \ker(\psi)$ est immédiat, parce que $\psi([2]P) = \psi(P)^2$. Par contre, l'autre inclusion apparaît presque toute seule dans les calculs de la démonstration cohomologique, alors qu'ici on n'obtient le résultat voulu qu'au prix de calculs venus de nulle part, et $\ker(\psi) \ni P = [2]Q$ surgit par magie. Pour vous éviter ça, croyez-moi, ce résultat est vrai.

Et on en arrive à un point crucial de la démonstration : soit S un ensemble fini de places, choisi de tel sorte que $\mathscr{O}_{K,S}$ soit principal (on peut, comme on l'a déjà signalé auparavant !), et que les $\alpha_i - \alpha_j$ soient des S-unités (c'est-à-dire des éléments de $\mathscr{O}_{K,S}^*$) ; ceci est possible en ajoutant à S les idéaux premiers $\mathfrak{p}$ tels que $\mathrm{ord}_{\mathfrak{p}}(\alpha_i - \alpha_j) \neq 0$.

À présent, j'écris x et y sous forme de fractions irréductibles A/B et C/D, avec $A, B, C, D \in \mathscr{O}_{K,S}$ [†]. Comme prévu, on obtient :

$$B^3C^2 = D^2 \prod_{i=1}^{3}(A - \alpha_i B).$$

Le S-entier D^2 étant premier avec C, il divise B^3. De même, B étant premier avec A, B^3 divise D^2. Alors, modulo un élément inversible, on a $B^3 = D^2$. Si on écrit $B = E^2$ et $D = E^3$, on se retrouve, pour résumer, avec :

$$C^2 = \prod_{i=1}^{3}(A - \alpha_i E^2). \tag{10}$$

De plus, les différents facteurs sont premiers entre eux : si p est un nombre premier dans $\mathscr{O}_{K,S}$ qui divise $A - \alpha_i E^2$ et $A - \alpha_j E^2$ (pour i et j distincts), alors p divise $(\alpha_i - \alpha_j)E^2$ et $(\alpha_i - \alpha_j)A$. Comme A et $E^2 = B$ sont premiers entre eux et $(\alpha_i - \alpha_j)$ est inversible, p n'a pas d'autre choix que d'être inversible également ! Ainsi, (10)

†. Je rappelle que le corps de fractions de $\mathscr{O}_K$ est K, donc c'est le cas également pour $\mathscr{O}_{K,S}$.

et le fait que les facteurs soient premiers entre eux impliquent que ce sont tous des carrés, à un élément inversible près. Bref ! $x(P) - \alpha_i = \epsilon_i t_i^2$, où ϵ_i peut être choisi dans $\mathscr{O}_{K,S}^*/\mathscr{O}_{K,S}^{*2}$, puisque les carrés n'importent pas pour ψ. On peut conclure :

$$\Gamma \simeq \psi(E(K)) \hookrightarrow (\mathscr{O}_{K,S}^*/\mathscr{O}_{K,S}^{*2})^3,$$

produit d'ensembles-quotient qui est fini par le théorème des unités généralisé, et le théorème de Mordell-Weil faible est démontré. □

EPIC

Je suis désolé de faire tomber la tension si brusquement après la démonstration du théorème de Mordell-Weil ; c'est juste pour signaler que le même genre de stratégie permet de montrer que le nombre de points entiers est, lui, fini, dans le cas où les coefficients de la courbe elliptique sont entiers. C'est le théorème de Siegel, dont on peut voir la démonstration dans [**Hd2**].

4. Pour aller plus loin

4.1. Théorème de Mordell-Weil : effectivité ?— Bien que le théorème de Mordell-Weil soit intéressant tel quel, son efficacité serait décuplée si la démonstration du théorème était effective. Il est en effet naturel de se demander, face à son énoncé, si on sait explicitement trouver le rang de la courbe elliptique, ainsi que les générateurs d'ordre fini ou non. Le cas des éléments d'ordre fini (qui forment le groupe de torsion, comme on dit dans la littérature) est plus simple à résoudre à la main, et est même complètement résolu pour plusieurs cas de courbes elliptiques (voir [**Hd3**]). On a par exemple le résultat suivant, dont la démonstration me semble hors de portée :

***Théorème 4.1* (Théorème de Mazur).** — *Le sous-groupe de torsion de $E(\mathbb{Q})$ est isomorphe soit à $\mathbb{Z}/m\mathbb{Z}$ avec $m \in [\![1, 10]\!] \cup \{12\}$, soit à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2m\mathbb{Z}$ avec $m \in [\![1, 4]\!]$.*

Alors, l'ordre des éléments de $E(\mathbb{Q})$ est soit entre 1 et 10, soit égal à 12.

Mais le cas général n'est pas résolu. Ceci pose problème parce que d'après la démonstration proposée, il faudrait savoir trouver des représentants de Γ. De $\Gamma \simeq \psi(E(K))$, on déduit que ceci revient à déterminer les points rationnels des courbes définies par $A - \alpha_i E^2 = \epsilon_i Z_i^2$ pour trois ϵ_i donnés de $(\mathscr{O}_{K,S}^*/\mathscr{O}_{K,S}^{*2})^3$. Il semble qu'un tel algorithme soit encore inconnu. On peut à la rigueur chercher des générateurs parce qu'on sait borner leurs hauteurs : On a vu que $\hat{h}$ est quadratique, et on sait montrer grâce au théorème de Minkowski qu'elle est définie positive [*] ; soit $\langle , \rangle$ son produit scalaire associé. Il est coutume de noter $\mathrm{Reg}(E(K)) = |\det(\langle P_i, P_j \rangle)|$ où les P_i sont des générateurs libres de $E(K)$.On a alors, par des résultats classiques de

∗. L'idée est belle, mais il faut y penser ! $\hat{h}$ induit une forme quadratique Q sur $\mathbb{R}^r$. On suppose que la forme quadratique est dégénérée, donc s'écrit $x_1^2 + \cdots + x_s^2$ avec $s < r$. Alors $\{\vec{x} \in \mathbb{R}^r | Q(\vec{x}) \leqslant \epsilon\}$ (cylindre de volume infini) contient un vecteur non nul d'après Minkowski, et ceci contredit le fait que seuls les points de torsion vérifient $\hat{h}(P) \leqslant \epsilon$ pour ϵ assez petit.

géométrie (dus à Hermite et à Hadamard), l'existence d'une base de P_i (modulo la torsion) tels que :

$$\mathrm{Reg}(E(K)) \leqslant \hat{h}(P_1)\cdots\hat{h}(P_r) \leqslant c^{r^2}\mathrm{Reg}(E(K)),$$

où P_1 est le point qui donne le minimum de $\hat{h}$ parmi les éléments non nuls du groupe libre.

4.2. Conjecture de Birch et Swinnerton-Dyer. — Le problème précédent aurait un élément de réponse avec la présente et difficile conjecture : la conjecture de Birch et Swinnerton-Dyer. Pour être en mesure de comprendre les termes de la conjecture, il est temps de présenter un des points de vue annoncés en introduction. Avant tout, parlons d'une des fonctions les plus importantes de l'arithmétique qui est, comme vous l'avez tous deviné :

$$\zeta(s) = \sum_{n\geqslant 1}\frac{1}{n^s} = \prod_{p \text{ premier}}\frac{1}{1-p^{-s}},$$

la fonction ζ de Riemann, pour $\Re(s) > 1$. La première égalité est la définition usuelle, la seconde égalité est une belle formulation synthétique de la décomposition unique en facteurs premiers dans $\mathbb{Z}$. Quelques réflexions sur ζ conduisent à une généralisation de cette fonction pour des courbes elliptiques E, et pour certaines courbes elliptiques, ζ_E a les mêmes zéros que $1 - a_pT + pT^2$ où p est un nombre premier, et $a_p = p + 1 - \mathrm{card}\ (E(\mathbb{F}_p))$, ces zéros étant de partie réelle $1/2$ (comme pour hypothèse de Riemann !). Ceci inspire la définition des fonctions L de Hasse-Weil :

Définition 4.2. — Si E est une courbe elliptique sur $\mathbb{Q}$, on pose $\mathrm{L}_E(s) = \prod_p L_p(s)$ où :

$$L_p(s) = \begin{cases} (1 - a_p p^{-s} + p^{1-2s})^{-1} & \text{si } E \text{ est une courbe elliptique modulo } p \\ \text{Des termes d'ajustement} & \text{s'il y a des singularités.} \end{cases}$$

Les termes d'ajustement dépendent des tangentes en la singularité, et ne nous intéressent pas ici, faute de place (ils ne concernent qu'un nombre fini de premiers p). Pour que ceci ait un sens, on doit pouvoir écrire E sous la forme

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

où les coefficients sont *entiers*, ce qui est toujours possible. Bref, L vérifie des choses très intéressantes (équation fonctionnelle, méromorphe sur un domaine qu'on voudrait large, produit infini...), mais dans le cadre de Mordell-Weil on a la conjecture suivante :

Conjecture 4.3 **(Conjecture de Birch et Swinnerton-Dyer)**
- *L'ordre d'annulation de* $\mathrm{L}_E(s)$ *en* $s = 1$ *égale* $r = \mathrm{rang}(E(\mathbb{Q}))$.
- *Si* $P_1, \ldots, P_r$ *forment une base de* $E(\mathbb{Q})$ *modulo la torsion, alors :*

$$\lim_{s\to 1}\frac{\mathrm{L}_E(s)}{(s-1)^r} = u\Omega_E \det(\langle P_i, P_j\rangle), \text{ où } u \in \mathbb{Q}^* \text{ et } \Omega_E = \int_{E(\mathbb{R})}\frac{\mathrm{d}x}{2y + a_1x + a_3}.$$

Le nombre rationnel u est plus explicite, et dépend d'un certain nombre de quantités : du cardinal du groupe de Shafarevitch-Tate, un groupe dont la finitude est encore discutée, et du cardinal d'un quotient de $E(\mathbb{Q}_p)$ ($\mathbb{Q}_p$ est la complétion p-adique de $\mathbb{Q}$) entre autres. Cette conjecture est assez complexe, puisqu'elle dépend d'autres conjectures (nullité de $\mathrm{L}_E(1)$ [(*)], finitude d'un groupe...). Cette conjecture est aussi fascinante, parce qu'à partir d'informations « locales » (réduction modulo p), on construit une fonction L qui donne une information « globale » sur la courbe, à savoir le rang. Mais il y a des raisons de penser qu'elle est vraie, puisqu'elle a été vérifiée sur certains cas particuliers, et parce que conformément à l'intuition $u\Omega_E \det(\langle P_i, P_j\rangle)$ est invariant par « morphismes de courbes elliptiques », ou plutôt *isogénies*. Enfin, il a été montré que $\mathrm{L}_E(1) = 0$ pour les courbes elliptiques dites de *type CM*. Ce problème n'est pas sous-estimé par l'Institut Clay qui a mis sa tête à prix : un million de dollars.

4.3. Lien avec les formes modulaires. — Puisque les courbes elliptiques sont essentiellement une branche de l'arithmétique, on peut se demander quel est le lien entre ces courbes et la cinquième opération de l'arithmétique, à savoir les formes modulaires (boutade souvent attribuée à Eichler) ! En effet, le lien entre fonctions L et courbes elliptiques ne s'arrête pas là, et n'est pas la seule façon d'étudier ces courbes.

Les formes modulaires peuvent être considérées comme faisant partie de l'arithmétique parce qu'elles sont de bons générateurs d'identités arithmétiques ; ceci mérite quelques définitions sur la question [(*)]. Je donne juste une définition :

Définition 4.4. — Une forme modulaire de poids k est une fonction méromorphe sur $\mathscr{H}$, ayant une limite finie quand $\Im(z) \to \infty$, et vérifiant, pour tous $z \in \mathscr{H}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$:

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^{-k} f(z). \tag{11}$$

Je rappelle que $\mathscr{H}$ est l'ensemble $\Im(z) > 0$, et qu'une matrice est dans $\mathrm{SL}_2(\mathbb{Z})$ pourvu que ses coefficients soient entiers, et que $ad - bc = 1$. Sous telle condition, f est 1-périodique, donc admet un développement de Fourier $f(z) = \sum_{n \geqslant 0} a_n q^n$ où $q = e^{2i\pi z}$. Il est parfois préférable de travailler avec d'autres groupes de matrices que $\mathrm{SL}_2(\mathbb{Z})$. Aussi note-t-on :

$$\Gamma(N) = \{M \in \mathrm{SL}_2(\mathbb{Z}) | M \equiv \mathrm{I}_2 \mod N\},$$

et on dit que f est modulaire de poids k de niveau N s'il vérifie (11) pour un groupe de matrices contenant $\Gamma(N)$. L'ensemble de ces f constitue $S_k(\Gamma(N))$.

∗. D'après la définition de L_E, on peut montrer la convergence pour $\Re(s) > 3/2$ au mieux. Pour s'assurer que $\mathrm{L}_E(1)$ est bien définie, on a besoin de l'équation fonctionnelle qui sera établie plus tard.

∗. les propositions intéressantes sur les formes modulaires peuvent se retrouver dans [**Ser**].

Pour comprendre l'énoncé « Toute courbe elliptique sur $\mathbb{Q}$ est modulaire », je commence par donner un exemple simple qui m'a été inspiré par [**D& S**] : considérons l'équation

$$Q : x^2 = d, \quad d \in \mathbb{Z}.$$

On peut, pour résoudre ceci, raisonner modulo p, et ceci conduit à la définition

$$a_p = \{\bar{x} \in \mathbb{F}_p | \bar{x}^2 = \bar{d}\} - 1,$$

qui vaut 1 ou -1. Le lemme chinois donne que si $n = \prod_p p^{\mathrm{ord}_p(n)}$, alors

$$a_n(Q) = \prod_p a_p(Q)^{\mathrm{ord}_p(n)}.$$

De $a_p(Q) = \left(\frac{d}{p}\right)$ (symbole de Legendre) pour $p \neq 2$ et la loi de réciprocité quadratique, on peut montrer que $\{a_2(Q), a_3(Q), \ldots\}$ est un ensemble de valeurs propres sur un $\mathbb{C}$-espace vectoriel associé à Q. Plus précisément :

$$\forall p, \forall n \in (\mathbb{Z}/4|d|\mathbb{Z})^*, T_p(f)(n) = a_p(Q) f(n),$$

où $f(n) = a_n(Q)$ pour tout $n \in (\mathbb{Z}/4|d|\mathbb{Z})^*$, et T_p est un endomorphisme de l'espace vectoriel des fonctions de $(\mathbb{Z}/4|d|\mathbb{Z})^*$ dans $\mathbb{C}$, défini par $T_p(f)(n) = f(pn)$ si p ne divise pas n, 0 sinon. Bref, $a_p(Q)$ est une valeur propre de l'opérateur T_p pour le vecteur propre f. On dit que f est fonction propre des T_p. Le cas s'étend aux formes modulaires en posant $a_p = p - \{(x, y) | (\bar{x}, \bar{y}) \in E(\mathbb{F}_p)\}$, et $f(\tau) = \sum_{n \geqslant 0} a_n(f) e^{2\pi i n \tau}$. Le théorème « modulaire » dit que toute courbe elliptique E est liée à une telle fonction f fonction propre des opérateurs de Hecke pour les valeurs propres a_p, qu'en plus $a_p(f) = a_p(E)$, et est une forme modulaire de poids 2 à un niveau N_E. Les grands résultats en ce sens sont :

***Théorème 4.5* (Théorème de Wiles).** — *Si $E(\mathbb{Q})$ est une courbe elliptique, la fonction L_E se prolonge analytiquement en une fonction entière qui vérifie, si on pose $\Lambda_E(s) = N_E^{s/2}(2\pi)^{-s}\Gamma(s)\mathrm{L}_E(s)$:*

$$\Lambda_E(s) = \pm\Lambda_E(2 - s)$$

L'énoncé de Wiles est en fait plus précis, puisqu'il déduit l'équation fonctionnelle de L d'une équation fonctionnelle pour la forme modulaire associée. Il implique également :

***Corollaire 4.6* (Théorème de Taniyama-Weil).** — *Soit E une courbe elliptique sur $\mathbb{Q}$, $\mathrm{L}_E(s) = \sum_{n \geqslant 1} a(n) n^{-s}$ sa fonction* L *et $f_E(\tau) = \sum_{n \geqslant 1} a(n) q^n$ la transformée de Mellin*(†) *inverse de $(2\pi)^{-s}\Gamma(s)\mathrm{L}_E(s)$. Alors f est dans $S_2(\Gamma_0(N))$ où N désigne le conducteur de E et f est une forme dite de Hecke.*

†. La transformée de Mellin est une opération proche de la transformée de Fourier ou de Laplace, définie par $Mf(s) = \int_0^\infty t^{s-1} f(t) \mathrm{d}t$, et est très utilisée en théorie analytique des nombres.

Reductio ad extremum, la puissance de ce corollaire se justifie par sa première application spectaculaire : si le théorème de Fermat avait une solution non triviale pour p premier, on pourrait construire une courbe elliptique *jument de Roland*, $y^2 = x(x-a^p)(x+b^p)$, dont la forme modulaire associée aurait des propriétés extraordinaires qui entraineraient, via des théorèmes et représentations galoisiennes qui ne sont pas de mon ressort, qu'elle est de niveau 2 et non nulle... Une telle forme n'existe pas !

5. Conclusion

Ainsi, alors que je pensais avoir tout vu des courbes elliptiques *via* les quelques conférences à leur sujet que j'ai pues voir, ma surprise vis-à-vis la richesse et la profondeur de la matière n'a fait que s'accroître au fil des jours, de même que la difficulté manifeste de ses problèmes. Mais comme le disait ln(3) dans l'*Iliade* :

> οἵσιν ἐπὶ Ζεὺς θῆκε κακὸν μόρον, ὡς καὶ ὀπίσσω Μάνθρώποισι πελώμεθ' ἀοίδιμοι ἐσσομένοισι. (‡)

Tout cela pour dire que la diversité des outils, aussi bien algébriques, analytiques et géométriques, réunis dans une sorte de *best-of* de tout ce que j'ai appris, m'a fasciné (et entre nous, c'est ce dont raffolent tous les amateurs d'arithmétique). La présence du dictionnaire trilingue voire n-lingue (qui traduit les courbes elliptiques en fonctions elliptiques, fonctions L, fonctions modulaires, variétés abéliennes...), et surtout la perspective d'y avoir beaucoup à apprendre (j'ai dû passer à côté de la géométrie algébrique, les surfaces de Riemann, la cohomologie galoisienne...) m'ont motivé et me motivent pour les années d'études à venir. D'ailleurs, le théorème de Mordell-Weil semble être un bon point d'appui pour une introduction à la théorie, puisque les démonstrations de ce théorème sont multiples et utilisent chacune un côté incontournable de la théorie. Pour compléter cette étude qui fut très enrichissante pour moi, je renvoie à [**Win2**].

THE END

To be continued... ?

‡. Zeus nous a fait un cruel destin, afin que nous soyons chantés des hommes à venir.

6. Annexe

*Démonstration du théorème 2.4 (*Nullstellensatz*).* — Pour ne pas alourdir la démonstration, je ne démontre pas qu'une K-algèbre qui est un corps de type fini est algébrique sur K, donc égale K si K est algébriquement clos. Une fois ce résultat (démontré dans [**Lng**]) admis, voici comment procéder : par hypothèse, les polynômes de $K[X_1, \dots, X_m, T]$ $P_1, \dots, P_m$ et $1 - TQ$ n'ont aucun zéro en commun dans K^{m+1}.

Alors, $(P_1, \dots, P_m, 1 - TQ)$ égale $K[X_1, \dots, X_m, T]$. En effet, dans le cas contraire, il existe un idéal maximal $\mathfrak{M}$ contenant cet idéal (l'anneau est noethérien, ou bien on utilise le lemme de Zorn), et le quotient de l'anneau par $\mathfrak{M}$ serait une extension algébrique de K, donc isomorphe à K. Je pourrais donc construire un zéro commun de K^{m+1} à tous les polynômes de $\mathfrak{M}$, en prenant $(x_1, \dots, x_m, t)$ où $x_i \equiv X_i \mod \mathfrak{M} \in K$, ce qui contredirait ce qu'on a dit un paragraphe plus haut.

Ainsi, $(P_1, \dots, P_m, 1 - TQ) = K[X_1, \dots, X_m, T]$, et en particulier

$$1 \in (P_1, \dots, P_m, 1 - TQ).$$

En substituant $T = 1/Q$, et en multipliant par Q^r pour r assez grand, on obtient le résultat voulu. □

Démonstration du corollaire 2.19 (somme de deux carrés). — Soit $C = B_f(\vec{0}, r)$. Elle est de mesure πr^2. Ses vecteurs entiers seraient de norme au carré x^2+y^2, et idéalement cette norme au carré devrait être un multiple de p, et suffisamment borné pour être exactement égale à p. Pour en faire un multiple de p, $x^2+y^2 \equiv 0 \mod p$ pour $y^2 \equiv -x^2 \mod p$. Une condition suffisante est $y \equiv ax \mod p$ avec $a^2 \equiv -1 \mod p$ [§].

Ceci étant dit, soit $\Lambda = \{(x, y) \in \mathbb{Z}^2 \mid y \equiv ax \mod p\} = (p, 0)\mathbb{Z} \oplus (1, a)\mathbb{Z}$ de déterminant p. Pour $r = 2\sqrt{\frac{p}{\pi}}$, $(x, y) \in \mathbb{Z}^2$ non nul vérifie $y \equiv ax \mod p$ comme désiré, donc $x^2 + y^2 \equiv 0 \mod p$ et $p | x^2 + y^2$. Or :

$$0 < x^2 + y^2 \leqslant \frac{4p}{\pi} < 2p,$$

donc $p = x^2 + y^2$.

EPIC

□

Ébauche de démonstration du théorème 2.27 (Bézout). — Définissons d'abord la multiplicité. Vu le contexte, il n'est pas absurde de travailler avec ce qu'on appelle les anneaux localisés en $P = (a, b) \in \mathbb{A}^2$, qui sont en fait les anneaux localisés en les $(X - a, Y - b)$:
$\mathscr{O}_P = \{R = \frac{P}{Q} \in K(X, Y) | Q(P) \neq 0\}$. En plus des propriétés propres aux anneaux locaux, on a $\mathscr{O}_P = K \oplus \underbrace{\ker(R \mapsto R(P))}_{=M_P}$ et $\mathscr{O}_P = K[x, y] + (f_1, f_2)_P$, où f_1 et f_2 sont les polynômes qui définissent respectivement $\mathscr{C}_1$ et $\mathscr{C}_2$, et $(f_1, f_2)_P$ est l'idéal dans $\mathscr{O}_P$ engendré par f_1 et f_2.

§. Qui existe précisément pour $p \equiv 1 \mod 4$. La raison simple, liée au fait que les carrés vérifient $x^{(p-1)/2} \equiv 1[p]$, est détaillée dans [**Per**] par exemple.

Ceci étant dit, j'appelle multiplicité de $\mathscr{C}_1$ et $\mathscr{C}_2$ en P la dimension du K-espace vectoriel $\mathscr{O}_P/(f_1,f_2)_P$. Elle correspond bien à l'idée qu'on pourrait avoir de la multiplicité d'un point d'intersection, puisque si $\mathscr{C}_1=\mathscr{C}\sqcup\mathscr{C}'$, alors $\mathrm{mult}(\mathscr{C}_1\cap\mathscr{C}_2,P)=\mathrm{mult}(\mathscr{C}\cap\mathscr{C}_2,P)+\mathrm{mult}(\mathscr{C}'\cap\mathscr{C}_2,P)$. Si $P\notin\mathscr{C}_1\cap\mathscr{C}_2$, alors $(f_1,f_2)_P=\mathscr{O}_P$ car l'idéal contient des inversibles, et la multiplicité est nulle. Par contre, si $P\in\mathscr{C}_1\cap\mathscr{C}_2$, $\mathrm{mult}(\mathscr{C}_1\cap\mathscr{C}_2,P)=1+\dim\left(M_P/(f_1,f_2)_P\right)$. Bref,

$$P\in\mathscr{C}_1\cap\mathscr{C}_2\Leftrightarrow\mathrm{mult}(\mathscr{C}_1\cap\mathscr{C}_2,P)\geqslant 1.$$

L'essentiel de la démonstration réside en la démonstration que $k[x,y]/(f_1,f_2)$ est de dimension d_1d_2. Alors, grâce au théorème de décomposition en modules localisés (on peut montrer que les modules localisés décrits ci-dessus sont les seuls), on aura le résultat, et ceci découle de raisonnements algébriques relativement basiques. □

Démonstration de la proposition 3.18 (presque-loi du parallélogramme)

Si P ou Q est nul, il n'y a rien à dire. Si $Q=\pm P$, on retrouve la proposition précédente. Je suppose donc qu'ils soient non nuls, et que $Q\neq\pm P$. Je remarque, pour commencer, que $h(P+Q)+h(P-Q)=h([1,x(P+Q)])+h([1,x(P-Q)])$. Or :

$$h([1,x(P+Q)])+h([1,x(P-Q)])=\ln(H([1,x(P+Q])H([1,x(P-Q)])).$$

L'égalité $\max(1,|\alpha|_v)\max(1,|\beta|_v)=\max(1,|\alpha+\beta|_v,|\alpha\beta|_v)$ pour v ultramétrique (facile à vérifier au cas par cas), et l'inégalité triangulaire pour les valeurs absolues archimédiennes donnent aisément que

$$\frac{1}{2}H(\alpha)H(\beta)\leqslant H(1,\alpha+\beta,\alpha\beta)\leqslant 2H(\alpha)H(\beta),\ \text{et} \tag{12}$$

$h(P+Q)+h(P-Q)=h(1,x(P+Q)+x(P-Q),x(P+Q)x(P-Q))+O(1)$, si bien que cette analyse nous conduit à introduire

$$\psi(P,Q)=(1,x(P)+x(Q),x(P)x(Q)),\mu(P,Q)=(P+Q,P-Q)$$

et, si je retiens l'attention sur les formules (4) et (5) qui s'intéressaient déjà à $P+Q$ et consorts, un choix de variable habile conduit au choix

$$\phi(T,U,V)=(U^2-4TV,2U(aT+V)+4bT^2,(aT-V)^2-4bTU),$$

qui fera apparaître le coefficient 2 bientôt [¶]. On a alors, vous le remarquez, $\psi\circ\mu=\phi\circ\psi$. Là encore, le calcul montre que $\phi(T,U,V)=(0,0,0)$ uniquement pour $(T,U,V)=(0,0,0)$, c'est-à-dire jamais, dans $\mathbb{P}^2(K)$. La proposition 3.15 s'applique alors immédiatement, pour fournir

$$h(\phi(T,U,V))=2h(T,U,V)+O(1).$$

Appliqué à notre bazar, ceci fournit :

$h(P+Q)+h(P-Q)=h(\psi\circ\mu(P,Q))+O(1)=2h(\psi(P,Q))+O(1)=2h(P)+2h(Q)+O(1)$,

toujours grâce à (12). □

¶. Dans l'idée : j'ai posé $U=x(P)+x(Q)$ et $V=x(P)x(Q)$.

Références

[B & C] Z.I. Borevitch et Igor Rostilavovich Shafarevich, *Théorie des nombres*, 245 pages, 1996.

[Bou] Nicolas Bourbaki, *Algèbre commutative, chapitres 5 à 7*, 351 pages, Masson, 1985.

[D& S] Fred Diamond et Jerry Shurman, *A first course in modular forms*, 436 pages, Springer, 2007.

[Har] Robin Hartshorne, *Algebraic Geometry*, 512 pages, Springer-Verlag, 1977.

[Hel] Yves Hellegouarch, *Invitation aux mathématiques de Fermat-Wiles*, 397 pages, Masson, 1997.

[Hd1] Marc Hindry, *Arithmétique*, 327 pages, Calvage et Mounet, 2008.

[Hd2] Marc Hindry, *Cours d'algèbre niveau L3 du magistère de l'ENS Cachan*, 89 pages, 2005.

[Hd3] Marc Hindry, *Why is it difficult to compute the Mordell-Weil group ?*, 17 pages, 2005.

[Jac] Nathan Jacobson, *Basic Algebra II*, 704 pages, Dover Publications, 2004.

[Lng] Serge Lang, *Algèbre*, 926 pages, Dunod, 2004.

[Per] Daniel Perrin, *Cours d'algèbre*, 208 pages, Ellipses, 1998.

[Rau] Gérard Rauch, *Les groupes finis et leurs représentations*, 144 pages, Ellipses, 2001.

[Sam] Pierre Samuel, *Théorie algébrique des nombres*, 132 pages, Hermann, 1997.

[S & T] Joseph H. Silvermann et John Tate, *Rational points on elliptic curves*, 291 pages, Springer, 1994.

[Ser] Jean-Pierre Serre, *Cours d'arithmétique*, 192 pages, PUF, 1994.

[Win] Bruno Winckler, *Recueil de blagues mathématiques et autres curiosités*, 115 pages, 2009.

[Win2] Bruno Winckler, *Les courbes elliptiques ; théorème de Mordell-Weil*, 39 pages, 2009.

Bruno Winckler