Chapitre I

Courbes elliptiques

Une courbe elliptique peut être définie comme une courbe projective lisse de degré 3 dans le plan projectif, munie d'un point origine; l'ensemble des points est alors muni d'une structure de groupe. La description la plus concrète provient du fait qu'on peut écrire leur équation affine sous la forme :

$$y^2 = x^3 + ax + b, \qquad avec \quad 4a^3 + 27b^2 \neq 0.$$

La théorie des courbes elliptiques est un merveilleux mélange de mathématiques élémentaires, profondes et sophistiquées, qui se situent de surcroît au croisement de multiples branches : arithmétique, géométrie algébrique, représentations de groupes, analyse complexe, etc. Nous donnons ici une introduction au sujet et démontrons les deux principaux théorèmes diophantiens : le groupe des points rationnels est un groupe de type fini (théorème de Mordell-Weil) et l'ensemble des points entiers est fini (théorème de Siegel). On évoque enfin le célèbre théorème de Wiles - qui aboutit à la preuve du théorème de Fermat - et la conjecture de Birch & Swinnerton-Dyer.

1. La loi de groupe sur une cubique

Ici le mot « cubique » désigne une courbe C dans le plan projectif $\mathbf{P}^2$ définie par une équation $F(X, Y, Z) = 0$ homogène de degré 3. La courbe est *lisse* si elle admet une tangente en chaque point, i.e. si $(\frac{\partial F}{\partial X}, \frac{\partial F}{\partial Y}, \frac{\partial F}{\partial Z}) \neq (0, 0, 0)$ (voir l'appendice B pour une introduction à la géométrie projective). Si $F \in K[X, Y, Z]$ on rappelle qu'on note $C(K)$ l'ensemble des points rationnels sur K c'est-à-dire l'ensemble $\{(x, y, z) \in \mathbf{P}^2(K) \mid F(x, y, z) = 0\}$.

1.1. Définition. Soit C une cubique lisse. Si P, Q sont des points distincts d'une cubique C, la droite joignant P, Q coupe la cubique en trois points P, Q et un troisième point R (éventuellement égal à P ou Q), on notera $R = P \circ Q$. Si $P = Q$ on fait la même opération avec la tangente en P à la courbe C. On définit une autre loi en choisissant un point $O \in C$, en posant $O' = O \circ O$ puis

$$P + Q := O \circ (P \circ Q) \qquad \text{et} \qquad -P := O' \circ P.$$

On dira que l'addition est définie par le procédé de « *tangentes et cordes* ».

1.2. Théorème. *La loi définie par le procédé de tangentes et cordes sur une cubique lisse est une loi de groupe commutatif dont l'élément neutre est O. Si $O \in C(K)$, alors $C(K)$ est un groupe abélien.*

Remarquons que, comme $P \circ Q = Q \circ P$, la loi $+$ est clairement commutative. On a bien $O + P = P$ puisque O, P et $O \circ P$ sont alignés. Soit $Q = -P$ alors les points Q, O' et P sont alignés donc $O' = Q \circ P$ et $O \circ O' = O$ donc $O = Q + P$. Le seul point délicat est de montrer l'associativité, pour cela nous aurons recours aux lemmes classsiques suivants.

1.3. Lemme. *Soient $P_1, \ldots, P_8 \in \mathbf{P}^2$ distincts ; supposons que 4 d'entre eux ne sont jamais alignés et 7 d'entre eux n'appartiennent jamais à une conique, alors l'espace vectoriel des polynômes homogènes de degré 3 s'annulant en $P_1, \ldots, P_8$ est de dimension 2.*

Démonstration. Notons n la dimension cherchée. Quelque soit les P_i, on a $n \geqslant 10 - 8 = 2$. Si, disons, P_1, P_2, P_3 sont alignés, choisissons P_9 sur la même droite d'équation $L = 0$. Toute cubique F s'annulant sur $P_1, \ldots, P_9$ est donc de la forme LQ avec Q s'annulant sur $P_4, \ldots, P_8$. Mais par cinq points dont quatre ne sont pas alignés, il ne passe qu'une conique disons $Q_0 = 0$ et F est un multiple de LQ_0. La dimension n_0 de l'espace de ces cubiques est donc égale à 1. Ainsi $n \leqslant n_0 + 1 = 2$. Supposons maintenant que $P_1, \ldots, P_6$ soit sur une conique $Q = 0$ et choisissons P_9 sur cette conique. Toute cubique F s'annulant sur $P_1, \ldots, P_9$ est donc de la forme LQ avec $L = 0$ équation de la droite (P_7, P_8). La dimension n_0 de l'espace de ces cubiques est donc égale à 1. Ainsi $n \leqslant n_0 + 1 = 2$. Dans le cas général (aucun triplet de points alignés, aucun sextuplet conconique), introduisons P_9, P_{10} situés sur la droite (P_1, P_2) d'équation $L = 0$. Si $n \geqslant 3$, il existera une cubique non triviale $F = 0$ passant par $P_1, \ldots, P_{10}$, mais alors $F = LQ$ et la conique d'équation $Q = 0$ passerait par $P_3, \ldots, P_8$. □

1.4. Lemme. *Soit $P_1, \ldots, P_9$ les points d'intersection de deux cubiques C_1 et C_2 dont l'une est irréductible. On suppose que $P_1, \ldots, P_8$ sont distincts. Si une cubique C passe par $P_1, \ldots, P_8$ alors elle passe par P_9.*

Démonstration. Supposons par exemple C_1 irréductible, alors elle ne peut contenir 4 points alignés ni 7 points conconiques. D'après le lemme précédent, l'espace vectoriel des cubiques s'annulant en $P_1, \ldots, P_8$ est de dimension 2 et est donc engendré par les équations de C_1 et C_2. □

Démonstration. (du théorème V-1.2) Soient P, Q, R trois points distincts de la cubique C la droite $L_1 = (P, Q)$ coupe C en P, Q, T, la droite $L_2 = (T, O)$ coupe C en T, O, T', la droite $L_3 = (R, T')$ coupe C en R, T', U et la droite $L_4 = (U, O)$ coupe C en U, O, U' de sorte que $(P+Q)+R = U'$. Par ailleurs la droite $M_1 = (Q, R)$ coupe C en Q, R, S, la droite $M_2 = (S, O)$ coupe C en S, O, S', la droite $M_3 = (P, S')$ coupe C en P, S', V et enfin la droite $M_4 = (V, O)$ coupe C en V, O, V' de sorte que $(Q + R) + P = V'$. On veut montrer que $U' = V'$, ce qui équivaut à $U = V$. Considérons pour cela les cubiques $C_1 = L_1 + M_2 + L_3$ et $C_2 = M_1 + L_2 + M_3$, on a alors

$$C \cap C_1 = \{P, Q, R, O, T, T', S, S', U\} \quad \text{et} \quad C \cap C_2 = \{P, Q, R, O, T, T', S, S', V\}.$$

Si les points P, Q, R, O, T, T', S, S' sont distincts, on peut conclure, par le lemme V-1.4, que $U = V$. C'est le cas en général et on conclut que l'égalité $(P+Q)+R = (Q+R)+P$ persiste en invoquant un argument de continuité (soit pour la topologie usuelle, si l'on travaille sur $\mathbf{R}$ ou $\mathbf{C}$, soit en utilisant la topologie de Zariski dans le cas général, voir l'appendice B, lemme **??**).□

figure

Remarquons que cette construction n'utilise que les deux cas les plus simples du théorème de Bézout : intersection avec une droite ou une conique.

Nous allons maintenant expliciter cette loi de groupe sur un modèle plus simple dit « de Weierstrass ».

1.5. Définition. Une cubique de Weierstrass est une courbe donnée par

une équation plane cubique de la forme :

$$Y^2Z = X^3 + aXZ^2 + bZ^3 \tag{I.1}$$

avec $\Delta := 4a^3 + 27b^2 \neq 0$.

1.6. Remarques. La condition $\Delta \neq 0$ signifie précisément que la courbe n'a pas de point singulier. La courbe définie par l'équation V-V.1 possède un point évident que l'on prend pour origine $O := (0, 1, 0)$ qui est un point d'inflexion, i.e. la tangente $Z = 0$ coupe la courbe uniquement en ce point et avec multiplicité 3. On peut montrer que toute cubique lisse possédant un point rationnel sur K est isomorphe à une cubique de Weierstrass, au moins lorsque K n'est pas de caractéristique 2 ou 3. Si l'on souhaite inclure le cas de caractéristique 2 et 3 (par exemple pour étudier les courbes elliptiques sur $\mathbf{F}_{2^f}$ ou $\mathbf{F}_{3^f}$), il faut prendre une équation un peu plus générale que V-V.1 de la forme :

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3. \tag{I.2}$$

En fait, cette dernière est l'équation générale d'une cubique ayant un point d'inflexion en $(0, 1, 0)$ avec tangente $Z = 0$. Notons que, si la caractéristique du corps est différente de 2 et 3 on peut aisément réduire l'équation (V-V.2) à la forme (V-V.1). En effet en posant $Y' := Y + (a_1X + a_3)/2$ on peut transformer l'équation en $Y'^2Z = X^3 + \frac{4a_2+a_1^2}{4}X^2Z + \dots$; en posant maintenant $X' := X + \frac{4a_2+a_1^2}{12}Z$ on trouvera une équation de Weierstrass du type (V-V.1).

En revenant au modèle de Weierstrass V-V.1, on travaillera souvent en coordonnées affines $x := X/Z$ et $y := Y/Z$ en considérant le point O comme « le point à l'infini ». L'équation affine est alors :

$$y^2 = x^3 + ax + b. \tag{I.3}$$

Un éventuel point singulier vérifierait $2y = 3x^2 + a = 0$ donc $y = 0$ et x racine double de $x^3 + ax + b = 0$ dont le discriminant est précisément $4a^3+27b^2$. Par ailleurs, si α est une racine de x^3+ax+b, le point $P := (\alpha, 0)$ est un point d'ordre 2, il y a donc trois points d'ordre 2.

1.7. Proposition. *(Loi de groupe explicite) Soit $P_1 = (x_1, y_1)$ et $P_2 = (x_2, y_2)$ des points de la courbe d'équation V-V.3 alors*

$$[-1](P) = (x_1, -y_1). \tag{I.4}$$

Si $P_2 = [-1](P_1)$ (i.e. si $x_1 = x_2$ et $y_2 = -y_1$) alors $P_1 + P_2 = O$. Si $P_2 = P_1$ posons $\lambda = \frac{3x_1^2+a}{2y_1}$ et $\mu = y_1 - \lambda x_1$, si $P_2 \neq \pm P_1$ (i.e. si $x_2 \neq x_1$)

posons $\lambda = \frac{y_1-y_2}{x_1-x_2}$ *et* $\mu = y_1 - \lambda x_1$, *alors :*

$$P_1 + P_2 = (\lambda^2 - x_1 - x_2, -\lambda^3 + \lambda(x_1 + x_2) - \mu) \tag{I.5}$$

Démonstration. Ecrivons $y = \lambda x + \mu$ l'équation de la droite (P_1, P_2) (resp. de la tangente à la courbe en P_1) si $P_1 \neq P_2$ (resp. si $P_1 = P_2$), alors λ et μ sont donnés comme dans l'énoncé. Si $P_3 = (x_3, y_3)$ est le troisième point d'intersection, on aura $P_1 + P_2 = (x_3, -y_3)$. Pour calculer les points d'intersection de la droite et de la courbe on substitue y pour obtenir

$$x^3 + ax + b - (\lambda x + \mu)^2 = x^3 - \lambda^2 x^2 + (a - 2\lambda\mu)x + (b - \mu^2) = 0$$

dont on connait déjà deux racines : x_1 et x_2, on en tire $x_1 + x_2 + x_3 = \lambda^2$ et $y_3 = \lambda x_3 + \mu$ d'où la proposition. □

Pour vérifier la continuité, au sens de la topologie de Zariski, de la loi d'addition, observons que :

$$\frac{y_1 - y_2}{x_1 - x_2} = \frac{x_1^2 + x_1 x_2 + x_2^2 + a}{y_1 + y_2}. \tag{I.6}$$

Pour usage futur, notons également les formules suivantes (que l'on vérifie par calcul direct) :

$$x(P+Q) + x(P-Q) = \frac{2(x(P) + x(Q))(a + x(P)x(Q)) + 4b}{(x(P) - x(Q))^2} \tag{I.7}$$

$$x(P+Q)x(P-Q) = \frac{(x(P)x(Q) - a)^2 - 4b(x(P) + x(Q))}{(x(P) - x(Q))^2} \tag{I.8}$$

$$x(2P) = \frac{x(P)^4 - 2ax(P)^2 - 8bx(P) + a^2}{4\,(x(P)^3 + ax(P) + b)}. \tag{I.9}$$

2. Hauteurs

2.1. Hauteurs de Weil

On introduit une notion précise de « taille » ou « complexité arithmétique » d'un point, que l'on baptisera hauteur. La première version s'appelle parfois hauteur de Weil et la version raffinée sur les courbes elliptiques, hauteur de Néron-Tate.

2.2. Définition. Soit P un point de $\mathbf{P}^n(\mathbf{Q})$, on peut choisir des coordonnées projectives $(x_0, \ldots, x_n)$ avec $x_i \in \mathbf{Z}$ et $\mathrm{pgcd}(x_0, \ldots, x_n) = 1$, on définit

alors la *hauteur* (resp. la *hauteur logarithmique*) de P par

$$H(P) := \max(|x_0|, \dots, |x_n|) \qquad (\text{resp. } h(P) := \log \max(|x_0|, \dots, |x_n|)).$$

Cette définition très simple et naturelle ne se transcrit pas si aisément aux coordonnées algébriques et il est plus commode techniquement de réinterpréter les hauteurs en terme de l'ensemble des valeurs absolues du corps.

2.3. Définition. Une *valeur absolue* v sur un corps K est une application $x \mapsto |x|_v$ de K vers $\mathbf{R}_+$ telle que pour tout $x, y \in K$,
i) $|x|_v = 0$ si et seulement si $x = 0$.
ii) $|xy|_v = |x|_v |y|_v$.
iii) $|x + y|_v \leqslant |x|_v + |y|_v$.
Si v vérifie l'inégalité plus forte $|x+y|_v \leqslant \max(|x|_v, |y|_v)$, on dira que v est *ultramétrique*.

2.4. Exemple. Les valeurs absolues standard sur $K = \mathbf{Q}$ sont, la valeur abolue usuelle notée $|x|$ ou $|x|_\infty$ et, pour chaque premier p, la valeur absolue p-adique définie par

$$|x|_p := p^{-\operatorname{ord}_p(x)}.$$

Les valeurs absolues p-adiques sont ultramétriques. On notera $M_{\mathbf{Q}}$ l'ensemble de ces valeurs absolues, que l'on appellera également *places* du corps $\mathbf{Q}$.

2.5. Théorème. *(Formule du produit pour* $\mathbf{Q}$*) Soit* $x \in \mathbf{Q}^*$ *alors*

$$\prod_{v \in M_{\mathbf{Q}}} |x|_v = 1. \tag{I.10}$$

Démonstration. Ecrivons $x = \pm p_1^{e_1} \dots p_r^{e_r}$ avec $e_i = \operatorname{ord}_{p_i}(x) \in \mathbf{Z}$. Pour $1 \leqslant i \leqslant r$, on a $|x|_{p_i} = p_i^{-e_i}$; si p n'apparaît pas dans x, on a $|x|_p = 1$ et la valeur absolue usuelle vaut $|x|_\infty = p_1^{e_1} \dots p_r^{e_r}$ d'où la formule. □

2.6. Corollaire. *Soit* $P \in \mathbf{P}^n(\mathbf{Q})$ *et* $(x_0, \dots, x_n)$ *des coordonnées projectives (quelconques) de* P*, on a*

$$H(P) = \prod_{v \in M_{\mathbf{Q}}} \max(|x_0|_v, \dots, |x_n|_v) \tag{I.11}$$

Démonstration. La formule du produit montre que le membre de droite est indépendant des coordonnées projectives. Si l'on choisit donc $x_i \in \mathbf{Z}$ premiers entre eux, on aura pour chaque p premier $\max(|x_0|_p, \dots, |x_n|_p) =$

1 et le membre de droite sera bien égal à $\max(|x_0|_\infty, \ldots, |x_n|_\infty)$, c'est-à-dire à $H(P)$. □

Pour généraliser les hauteurs aux points à coordonnéees algébriques, nous allons définir les valeurs absolues standard sur un corps de nombres K.

2.7. Exemple. Soit K un corps de nombres, avec r_1 plongement réels et r_2 paires de plongements complexes de sorte que $n := [K : \mathbf{Q}] = r_2 + 2r_2$. Chaque plongement $\sigma : K \hookrightarrow \mathbf{R}$ ou $\mathbf{C}$ produit une valeur absolue en composant avec le module; si le plongement est complexe, σ et son conjugué donneront la même valeur absolue, on dispose donc de $r_1 + r_2$ valeurs absolues, dont on note l'ensemble $M_{K,\infty}$:

$$|x|_\sigma := \begin{cases} |\sigma(x)| & \text{pour } \sigma \text{ réel} \\ |\sigma(x)|^2 & \text{pour } \sigma \text{ complexe.} \end{cases}$$

Si maintenant p premier se factorise en $p\mathscr{O}_K = \wp_1^{e_1} \ldots \wp_g^{e_g}$ avec $\mathrm{N}\,\wp_i = p^{f_i}$ et $\sum_{i=1}^g e_i f_i = n$, on peut définir pour chaque idéal premier $\wp$ une valeur absolue

$$|x|_\wp := \mathrm{N}\,\wp^{-\operatorname{ord}_\wp(x)}.$$

Il découle de ces choix que, pour $x \in K$ on a :

$$\prod_{\wp \mid p} |x|_\wp = \left|\mathrm{N}_{\mathbf{Q}}^K(x)\right|_p \qquad \text{et} \qquad \prod_{v \in M_{K,\infty}} |x|_v = \left|\mathrm{N}_{\mathbf{Q}}^K(x)\right|_\infty \tag{I.12}$$

En effet , si $x\mathscr{O}_K = \prod_\wp \wp^{\operatorname{ord}_\wp(x)}$ on peut écrire

$$\pm\,\mathrm{N}_{\mathbf{Q}}^K(x) = \mathrm{N}\,(x\mathscr{O}_K) = \prod_\wp \mathrm{N}\,\wp^{\operatorname{ord}_\wp(x)} = \prod_p p^{\sum_{\wp \mid p} f_\wp \operatorname{ord}_\wp(x)}.$$

On obtient donc

$$\left|\mathrm{N}_{\mathbf{Q}}^K(x)\right|_p = p^{-\sum_{\wp \mid p} f_\wp \operatorname{ord}_\wp(x)} = \prod_{\wp \mid p} \mathrm{N}\,\wp^{-\operatorname{ord}_\wp(x)} = \prod_{\wp \mid p} |x|_\wp$$

Pour les places archimédiennes on obtient

$$\left|\mathrm{N}_{\mathbf{Q}}^K(x)\right| = \left|\prod_{\sigma : K \hookrightarrow \mathbf{C}} \sigma(x)\right| = \prod_{i=1}^{r_1} |\sigma_i(x)| \prod_{j=1}^{r_2} \left|\sigma_j(x)\overline{\sigma_j(x)}\right| = \prod_{v \in M_{K,\infty}} |x|_v\,.$$

On a alors la formule analogue au théorème V-2.5 :

2.8. Théorème. *(Formule du produit pour K) Soit $x \in K^*$ alors*

$$\prod_{v \in M_K} |x|_v = 1. \tag{I.13}$$

Démonstration. On regroupe les places de K en places au dessus d'une certaine place de $\mathbf{Q}$ et on utilise la formule précédente :

$$\prod_{w \in M_K} |x|_w = \prod_{v \in M_{\mathbf{Q}}} \prod_{w \mid v} |x|_w = \prod_{w \in M_{\mathbf{Q}}} \left|\mathrm{N}_{\mathbf{Q}}^K x\right|_v = 1. \qquad \square$$

2.9. Définition. Soit $P \in \mathbf{P}^n(K)$ et $(x_0, \ldots, x_n)$ des coordonnées projectives (quelconques) de P, la *hauteur* relative au corps K est le nombre

$$H_K(P) = \prod_{v \in M_K} \max\left(|x_0|_v, \ldots, |x_n|_v\right). \tag{I.14}$$

Si $\alpha \in K$ on définit la *hauteur* de α (relative au corps K) comme la hauteur du point $(1, \alpha) \in \mathbf{P}^1(K)$.

Le lien entre la hauteur d'un nombre algébrique et son polynôme minimal est le suivant.

2.10. Lemme. *Soit α un nombre algébrique, écrivons son polynôme minimal dans $\mathbf{Z}[X]$ sous la forme $P(X) = a_0(X-\alpha_1)\ldots(X-\alpha_d) = a_0X^d + \ldots$ et $K = \mathbf{Q}(\alpha)$, alors :*

$$H_K(\alpha) = |a_0| \prod_{i=1}^{d} \max\{1, |\alpha_i|\} \tag{I.15}$$

Démonstration. Tout d'abord on a $\prod_{v \in M_{K,\infty}} \max(1, |\alpha|_v) = \prod_{i=1}^{d} \max\{1, |\alpha_i|\}$. Ensuite en considérant l'équation $a_0\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0$ on voit que :

$$\prod_{\wp} \max(1, |\alpha|_\wp) = \prod_{\wp} \mathrm{N}\,\wp^{\max\{0, -\operatorname{ord}_\wp(\alpha)\}} = |a_0|. \qquad \square$$

La hauteur d'un point considéré dans diverses extensions varie simplement.

2.11. Lemme. *Soit K' une extension finie de K et $P \in \mathbf{P}^n(K)$ alors*

$$H_{K'}(P) = H_K(P)^{[K':K]}. \tag{I.16}$$

Démonstration. Soit $(x_0, \ldots, x_n)$ des coordonnées projective de P, on peut supposer $x_i \in K$. Si v est une place de K et w parcourt les places de K'

au dessus de K, on a clairement

$$\prod_{w\,|\,v} \max|x_i|_w = \prod_{w\,|\,v} \max|x_i|_v^{e_w f_w} = \max|x_i|_v^{[K':K]}.$$

Donc on a bien :

$$\begin{aligned} H_{K'}(P) &= \prod_{w\in M_{K'}} \max|x_i|_w = \prod_{v\in M_K}\prod_{w\,|\,v} \max|x_i|_w \\ &= \prod_{v\in M_K} \max|x_i|_v^{[K':K]} = H_K(P)^{[K':K]}. \end{aligned}$$

□

Ce lemme permet de définir la hauteur absolue.

2.12. Définition. On définit $H : \mathbf{P}^n(\bar{\mathbf{Q}}) \to \mathbf{R}$ ainsi : si $P \in \mathbf{P}^n(K)$, on pose

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

Le principal mérite de cette fonction hauteur est le théorème de finitude suivant dont la première partie est due à Northcott et la seconde à Kronecker.

2.13. Théorème. *(Northcott, Kronecker) Soit $d \geqslant 1$ et $X > 0$ alors l'ensemble $S(n,d,X) = \{P \in \mathbf{P}^n(\bar{\mathbf{Q}}) \mid [\mathbf{Q}(P) : \mathbf{Q}] \leqslant d,\ H(P) \leqslant X\}$ est fini. De plus on a $H(P) > 1$ sauf si le point P possède des coordonnées projectives égales à zéro ou une racine de l'unité.*

Démonstration. Soit $P = (x_0, \ldots, x_n) \in \mathbf{P}^n(\bar{\mathbf{Q}})$, quitte à permuter les coordonnées, on peut supposer $x_0 \neq 0$, alors on peut écrire $P = (1, \alpha_1, \ldots, \alpha_n)$ avec α_i algébrique. On a trivialement que $H(\alpha_i) \leqslant H(P)$ et $[\mathbf{Q}(\alpha_i) : \mathbf{Q}] \leqslant [\mathbf{Q}(P) : \mathbf{Q}]$ donc il suffit de montrer que l'ensemble de nombres algébriques $\{\alpha \in \bar{\mathbf{Q}} \mid [\mathbf{Q}(\alpha) : \mathbf{Q}] \leqslant d,\ H(\alpha) \leqslant X\}$ est fini. Or une borne sur le degré et la hauteur donne, d'après le lemme V-2.10 une borne pour les coefficients du polynôme minimal de α, ce qui démontre la finitude. Pour la dernière assertion, on peut de nouveau ne considérer que $P = (1, \alpha_1, \ldots, \alpha_n)$ avec α_i algébrique. Si $H(P) = 1$ alors $|\alpha_i|_v \leqslant 1$ (pour tout i et tout v) donc cela reste vrai pour α^m ; ainsi l'ensemble des points $(1, \alpha_1^m, \ldots, \alpha_n^m)$ est fini, ce qui entraîne que chaque α_i est nul ou est une racine de l'unité. □

2.14. Lemme. *Soient α, β deux nombres algébriques alors*

$$\frac{1}{2} H(\alpha)H(\beta) \leqslant H(1, \alpha+\beta, \alpha\beta) \leqslant 2H(\alpha)H(\beta) \tag{I.17}$$

Démonstration. On peut raisonner cas par cas suivant que $|\alpha|_v \leqslant |\beta|_v \leqslant 1$

ou $|\alpha|_v \leqslant 1 \leqslant |\beta|_v$ ou $1 \leqslant |\alpha|_v \leqslant |\beta|_v$ Lorsque v est une valeur absolue ultramétrique, on vérifie directement l'égalité

$$\max(1, |\alpha+\beta|_v, |\alpha\beta|_v) = \max(1, |\alpha|_v)\max(1, |\beta|_v)\,.$$

Pour une valeur absolue archimédienne, on obtient un encadrement

$$\frac{1}{2}\max(1, |\alpha|_v)\max(1, |\beta|_v) \leqslant \max(1, |\alpha+\beta|_v, |\alpha\beta|_v) \leqslant 2\max(1, |\alpha|_v)\max(1, |\beta|_v)\,.$$

En faisant le produit de ces inégalités on obtient le lemme. □

2.15. Théorème. *Soit $P_0, \dots, P_m$ une famille de polynômes homogènes de degré d en $x = (x_0, \dots, x_n)$. Soit Z le lieu des zéros communs des P_i et $\Phi : \mathbf{P}^n \setminus Z \to \mathbf{P}^m$ l'application définie par $\Phi(x) = (P_0(x), \dots, P_m(x))$.*

i) Il existe une constante $C_1 = C_1(\Phi)$ telle que pour $x \in (\mathbf{P}^n \setminus Z)(\bar{\mathbf{Q}})$ on ait

$$H(\Phi(x)) \leqslant C_1 H(x)^d. \tag{I.18}$$

ii) Soit V une sous-variété fermée de $\mathbf{P}^n$ telle que $V \cap Z = \emptyset$, alors, il existe deux constantes $C_1 = C_1(\Phi)$ et $C_2 = C_2(\Phi)$ telles que pour $x \in V(\bar{\mathbf{Q}})$ on ait

$$C_2 H(x)^d \leqslant H(\Phi(x)) \leqslant C_1 H(x)^d. \tag{I.19}$$

Démonstration. La première inégalité se déduit d'une application répétée de l'inégalité triangulaire (usuelle et ultramétrique). Notons $x = (x_0, \dots, x_n)$ et $x^i := x_0^{i_0} \dots x_n^{i_n}$ et appelons K un corps de rationalité de x. Écrivons $P_i = \sum_j a_j^{(i)} x^j$ et notons $N = \binom{n+d}{d}$ le nombre de monomes de degré d et posons $N_v = 1$ pour les places ultramétriques et $N_v = N$ pour les places archimédiennes. On peut alors écrire, pour chaque place v de K :

$$|P_i(x)|_v \leqslant N_v \max_j \left|a_j^{(i)}\right|_v \max_i |x_i|_v^d\,.$$

En posant $A_v = \max_{i,j} \left|a_j^{(i)}\right|_v$, on voit que $A_v = 1$ sauf pour un nombre fini de places. On a donc

$$H_K(\Phi(x)) = \prod_v \max_i |P_i(x)|_v \leqslant \prod_v N_v A_v \max_i |x_i|_v^d = \left(\prod_v N_v A_v\right) H_K(x)^d$$

et, en prenant les racines $[K : \mathbf{Q}]$-èmes, on obtient la première inégalité avec $C := (\prod_v N_v A_v)^{1/[K:\mathbf{Q}]}$. Pour la deuxième inégalité, on a recours au théorème des zéros de Hilbert (voir théorème **??**) qui affirme, au vu des hypothèses que, si $Q_1 = \dots = Q_r = 0$ est un système d'équations de V, il existe des polynômes $A_i^{(j)}$ et $B_i^{(j)}$ et un entier $M \geqslant 1$ tels que

$$X_j^M = \sum_{i=0}^{m} A_i^{(j)} P_i + \sum_{i=1}^{r} B_i^{(j)} Q_i.$$

Notons de plus que l'on peut supposer les $A_i^{(j)}$ homogènes de degré $M-d$. Quand on l'applique à un point $x \in V$ on obtient

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

En appliquant comme précédemment l'inégalité triangulaire et en notant $n_v = 1$ (resp. $n_v = m+1$) si v est ultramétrique (resp. v est archimédienne), on obtient

$$|x_j|_v^M \leqslant n_v \max_i \left|A_i^{(j)}(x)\right|_v \max_i |P_i(x)|_v \leqslant A'_v \max_i |x_i|_v^{M-d} \max_i |P_i(x)|_v$$

avec $A'_v = 1$ sauf pour un nombre fini de places. On obtient alors

$$\max_j |x_j|_v^d \leqslant A'_v \max_i |P_i(x)|_v .$$

En faisant le produit sur les places v et en prenant la racine $[K : \mathbf{Q}]$-ème on obtient le résultat escompté. □

Notation. On posera $h_K = \log H_K$ et $h = \log H$ et on l'appellera *hauteur logarithmique*. Avec cette convention, la conclusion des inégalités ii) du théorème précédent peuvent se réécrire

$$h(\Phi(x)) = dh(x) + O(1).$$

Revenons maintenant aux courbes elliptiques et définissons une hauteur de Weil.

2.16. Définition. Soit $E \subset \mathbf{P}^2$ une courbe elliptique donnée par une équation de Weierstrass $Y^2Z = X^3 + aXZ^2 + bZ^3$, pour $P \in E(\bar{Q})$, on définit la *hauteur*[1] de P :

$$h(P) = \begin{cases} h(x(P)) & \text{si } P \neq 0_E \\ 0 & \text{si } P = 0_E. \end{cases}$$

2.17. Théorème. *La hauteur sur E est symétrique (i.e. on a $h(-P) = h(P)$) et vérifie presque la loi du parallélogramme, c'est-à-dire :*

$$h(P+Q) + h(P-Q) = 2h(P) + 2h(Q) + O(1). \tag{I.20}$$

Démonstration. La formule est trivialement vraie si P ou Q est nul, on peut donc supposer que $P, Q \in E \setminus \{0_E\}$. Posons $x_1 = x(P)$, $x_2 = x(Q)$, $x_3 = x(P+Q)$ et $x_4 = x(P-Q)$, puis $x_1 + x_2 = u$, $x_1x_2 = v$. La formule

[1] Il s'agit de hauteur logarithmique ; par ailleurs, pour des raisons sans importance ici cette hauteur est égale 2 fois la hauteur plus utilisée.

(V.7) se réécrit :

$$\begin{cases} x_3 + x_4 &= \dfrac{2u(a+v)+4b}{u^2-4v} \\ x_3x_4 &= \dfrac{(v-a)^2-4bu}{u^2-4v} \end{cases}$$

Ainsi, si l'on introduit l'application de $\mathbf{P}^2$ dans $\mathbf{P}^2$:

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

on remarque que les trois polynômes n'ont pas de zéros communs dans $\mathbf{P}^2$ (la vérification, qui utilise la condition $4a^3 + 27b^2 \neq 0$ est laissée en exercice). D'après la deuxième partie du théorème V-2.15 on obtient donc que

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

Par ailleurs, si on pose $\psi : (E \setminus \{0_E\})^2 \to \mathbf{P}^2$ définie par $\psi(P,Q) = (1, x(P)+x(Q), x(P)x(Q))$ et $\mu(P,Q) = (P+Q, P-Q)$ on voit que, d'une part

$$h(\psi(P,Q) = h(x(P)) + h(x(Q)) + O(1)$$

d'après le lemme V-2.14, et que d'autre part, en invoquant cette fois les formules V.7, on a :

$$\psi \circ \mu = \Phi \circ \psi.$$

On obtient ainsi

$$\begin{aligned} h(P+Q) + h(P-Q) &= h(x(P+Q)) + h(x(P-Q)) \\ &= h(1, x(P+Q)+x(P-Q), x(P+Q)x(P-Q)) + O(1) \\ &= h(\psi \circ \mu(P,Q)) + O(1) \\ &= h(\Phi(\psi(P,Q))) + O(1) \\ &= 2h(\psi(P,Q)) + O(1) \\ &= 2h(P) + 2h(Q) + O(1). \end{aligned}$$

□

2.18. Corollaire. *Il existe une constante (dépendant de E) telle que la hauteur sur E vérifie :*

$$-C_1 \leqslant h([2](P)) - 4h(P) \leqslant C_1 \tag{I.21}$$

Démonstration. Il suffit de prendre $P = Q$ dans l'énoncé précédent. □

2.19. Hauteurs de Néron-Tate

2.20. Lemme. *Soit $h : S \to \mathbf{R}$ et $f : S \to S$ telles que $|h \circ f - dh| \leqslant C$ sur S avec $d > 1$, alors pour tout $s \in S$ la suite $d^{-n}h(f^n(x))$ est une suite*

convergente de limite que l'on notera $\hat{h}_f(x)$ *et qui vérifie*

$$\left|h(x)-\hat{h}_f(x)\right| \leqslant \frac{C}{d-1} \tag{I.22}$$

$$\hat{h}_f(f(x)) = d\hat{h}_f(x) \tag{I.23}$$

Démonstration. En écrivant l'inégalité de l'énoncé au point f^{k-1} et en divisant par d^k, on obtient

$$-\frac{C}{d^k} \leqslant d^{-k}h(f^k(x)) - d^{-k+1}h(f^{k-1}(x)) \leqslant \frac{C}{d^k}.$$

En sommant ces inégalités entre n et m (avec disons $n<m$) on en tire

$$-\frac{C}{d^n(d-1)} \leqslant d^{-m}h(f^m(x)) - d^{-n}h(f^n(x)) \leqslant \frac{C}{d^n(d-1)}.$$

Ainsi $d^{-n}h(f^n(x))$ est bien une suite de Cauchy, et on peut noter $\hat{h}_f(x)$ sa limite. En faisant tendre m vers l'infini, on obtient alors

$$-\frac{C}{d^n(d-1)} \leqslant \hat{h}_f(x) - d^{-n}h(f^n(x)) \leqslant \frac{C}{d^n(d-1)},$$

et en particulier $\frac{C}{d-1} \leqslant \hat{h}_f(x) - h(x) \leqslant \frac{C}{d-1}$. Enfin on peut calculer

$$\hat{h}_f(f(x)) = \lim_{n\to\infty} d^{-n}h(f^{n+1}(x)) = d\lim_{n\to\infty} d^{-n-1}h(f^{n+1}(x)) = d\hat{h}_f(x). \quad \square$$

En appliquant ce lemme à la hauteur de Weil d'une courbe elliptique et au morphisme $[2] : E \to E$ (avec $d=4$) on obtient

2.21. Théorème. *(Néron-Tate) Soit E une courbe elliptique définie sur un corps de nombres K. On peut définir une hauteur dite « canonique » ou « de Néron-Tate » par la formule*

$$\hat{h}(P) = \lim_{n\to\infty} \frac{h(x(2^nP))}{4^n} \tag{I.24}$$

Cette hauteur sur E vérifie la loi du parallélogramme :

$$\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q) \tag{I.25}$$

et est donc quadratique. En particulier $\hat{h}(mP) = m^2\hat{h}(P)$ et $\hat{h}(P) = 0$ si et seulement si P est un point de torsion. De plus on a $\hat{h}(P) = h(P) + O(1)$.

Démonstration. On applique le lemme V-2.20 à la hauteur $h(P) = h(x(P))$ et à l'application $P \mapsto [2](P)$ avec $d=4$. L'inégalité du théorème V-2.17 appliquée aux points $P' = [2^n](P)$ et $Q' = [2^n](Q)$ donne

$$-\frac{C}{4^n} \leqslant 4^{-n}h([2^n](P+Q)) + h([2^n](P-Q)) - 2h([2^n](P)) - 2h([2^n](Q) \leqslant \frac{C}{4^n},$$

qui, en faisant tendre n vers l'infini donne la formule cherchée. Ainsi $\hat{h}$ est quadratique et vérifie en particulier $\hat{h}(mP) = m^2\hat{h}(P)$. Si $mP = 0$ on en tire immédiatement que $\hat{h}(P) = 0$. Inversement si $\hat{h}(P) = 0$ alors pour tout $m \in \mathbf{Z}$, on a $\hat{h}(mP) = 0$ donc l'ensemble $\{mP \mid m \in \mathbf{Z}\}$ est de hauteur bornée donc fini, ce qui entraîne que P est de torsion. □

Le dernier résultat peut s'interpréter en disant que la forme quadratique sur le réseau $E(K)/E(K)_{\text{torsion}}$ est non dégénérée. On peut être un peu plus précis et montrer le théorème suivant.

2.22. Théorème. *La forme quadratique réelle $E(K) \otimes \mathbf{R} \to \mathbf{R}$ induite par $\hat{h}$ est définie positive.*

Remarquons que le fait qu'une forme quadratique $Q(x)$ vérifie que, pour tout $x \in \mathbf{Q}^n \setminus \{0\}$, on a $Q(x) > 0$, n'entraîne pas qu'elle soit *définie* positive (penser à $Q(x_1, x_2) = (x_1 + x_2\sqrt{2})^2$).

Démonstration. Notons Q la forme quadratique sur $\mathbf{R}^n$ déduite de $\hat{h}$ par tensorisation avec $\mathbf{R}$. Elle est clairement positive, supposons qu'elle soit dégénérée, alors elle s'écrit après changement de base $Q(x_1, \ldots, x_r) = x_1^2 + \cdots + x_s^2$ avec $s < r$. On en tire facilement que les ensembles $\{x \in \mathbf{R}^r \,; Q(x) \leqslant \epsilon\}$ sont des cylindres de volume infini pour tout $\epsilon > 0$ et qui contiennent donc un point non nul de tout réseau d'après le théorème de Minkowski (**??**). Ceci contredit le fait que l'ensemble $\{P \in E(K) \mid \hat{h}(P) \leqslant \epsilon\}$ est réduit au sous-groupe de torsion pour ϵ assez petit. □

Ainsi l'on dispose d'un *produit scalaire* sur $E(K) \otimes \mathbf{R}$ qu'on peut définir par

$$\langle P, Q\rangle := \frac{1}{2}\left(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)\right).$$

En anticipant sur le théorème démontré au paragraphe suivant (le groupe $E(K)$ est de type fini), on peut tenter de préciser la taille des générateurs de $E(K)$ de la façon suivante.

2.23. Définition. Soit $P_1, \ldots, P_r$ une base de $E(K)$ modulo le sous-groupe fini de torsion F, on définit le *régulateur* de E comme :

$$\mathrm{Reg}(E/K) := \det\left(\langle P_i, P_j\rangle\right)_{1\leqslant i,j\leqslant r}.$$

et on définit la hauteur minimale d'un point d'ordre infini :

$$\hat{h}_{\min}(E/K) := \min_{P\in E(K)\setminus F} \hat{h}(P).$$

Ces deux quantités sont exactement les quantités nécessaires pour borner

la hauteurs de générateurs du groupe de Mordell-Weil $E(K)$ en vertu du résultat suivant de géométrie des nombres, dû à Hermite (voir exercice **??**).

2.24. Proposition. *Il existe des constantes C_n telles que, si L est un réseau de $\mathbf{R}^n$ muni de la norme euclidienne, alors il existe $e_1, \dots, e_n$ base de L telle que*

$$\det(L) \leqslant \|e_1\| \dots \|e_n\| \leqslant C_n \det(L).$$

3. Le théorème de Mordell-Weil

Le but de ce paragraphe est de démontrer le théorème suivant.

3.1. Théorème. *(Mordell-Weil) Soit E une courbe elliptique définie sur un corps de nombres K (par exemple $K = \mathbf{Q}$), le groupe $E(K)$ est un groupe de type fini.*

Une étape intermédiaire importante est le résultat suivant.

3.2. Théorème. *(Mordell-Weil « faible ») Soit E une courbe elliptique définie sur un corps de nombres K (par exemple $K = \mathbf{Q}$), le groupe $E(K)/2E(K)$ est fini.*

En fait le théorème de Mordell-Weil « faible » joint à la théorie des hauteurs du paragraphe précédent entraîne le théorème V-3.1 grâce au lemme de descente suivant.

3.3. Lemme. *Soit G un groupe muni d'une forme quadratique $q : G \to \mathbf{R}$ tel que les ensembles $\{x \in G \mid q(x) \leqslant X\}$ soient finis pour tout $X \in \mathbf{R}$ et tel que $G/2G$ soit fini. Le groupe G est de type fini. Plus précisément, si S est un ensemble de représentants modulo $2G$ et si $C := \max_{x \in S} q(x)$ alors $\{x \in G \mid q(x) \leqslant C\}$ engendre G.*

Démonstration. Notons $|x| := \sqrt{q(x)}$ de sorte que $|mx| = m|x|$ et $|x+y| \leqslant |x| + |y|$ (pour $x, y \in G$ et $m \in \mathbf{N}$). Soit $x \in G$ avec $q(x) > C$, on peut définir une suite x_n de points de G ainsi : on part de $x_0 = x$ puis on écrit $x_0 = y_1 + 2x_1$ avec $y_1 \in S$ et $x_1 \in G$, puis $x_1 = y_2 + 2y_2$ etc. On observe que

$$|x_1| = \left|\frac{x_0 - y_1}{2}\right| \leqslant \left(\frac{|x_0| + |y_1|}{2}\right) \leqslant \left(\frac{|x_0| + \sqrt{C}}{2}\right) < |x_0|$$

On peut itérer ce procédé et obtenir une suite vérifiant $|x_n| < |x_{n-1}| < \dots < |x_1| < |x_0|$ tant que $|x_n| > \sqrt{C}$. L'hypothèse de finitude entraîne que,

au bout d'un nombre fini d'étapes, on aura $|x_n| \leqslant \sqrt{C}$. Le point $x = x_0$ s'exprime comme combinaison des y_i et de x_n qui sont tous dans l'ensemble fini $\{y \in G \mid q(y) \leqslant C\}$ qui engendre donc bien G. □

Donnons maintenant le plan de la démonstration du théorème V-3.2. Nous supposerons pour simplifier que l'équation de la courbe s'écrit :

$$y^2 = f(x) = x^3 + ax + b = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3),$$

c'est-à-dire que nous supposerons les racines de f sont rationnelles sur K. En particulier on notera les points de 2-torsion $P_i = (\alpha_i, 0)$ sont dans $E(K)$. Cela ne restreint pas la généralité de la démonstration du théorème de Mordell-Weil puisqu'on peut toujours remplacer K par $K(\alpha_1, \alpha_2, \alpha_3)$; cependant, d'un point de vue algorithmique, il vaut mieux rester sur K, nous indiquerons en fin de paragraphe comment modifier la preuve.

3.4. Définition. On définit une application $\psi = (\psi_1, \psi_2, \psi_3) : E(K) \to \left(K^*/K^{*2}\right)^3$ par les formules suivantes :

$$\psi_i(P) = \begin{cases} x(P) - \alpha_i & \text{si } P \neq P_i, O_E \\ (\alpha_i - \alpha_j)(\alpha_i - \alpha_k) & \text{si } P = P_i \\ 1 & \text{si } P = 0_E. \end{cases}$$

3.5. Remarque. Dans l'optique de définir un *homomorphisme* ψ, la formule pour $P = P_i = (\alpha_i, 0)$ est naturelle car $(x - \alpha_i)K^{*2} = (x - \alpha_j)(x - \alpha_j)K^{*2}$; une autre définition possible serait de prendre $\psi_i(P_i) = f'(\alpha_i) \bmod K^{*2}$.

On démontre alors les trois lemmes suivants qui achèvent la démonstration du théorème V-3.2, puisque $E(K)/2E(K) \cong \psi(E(K))$.

3.6. Lemme. *L'application $\psi : E(K) \to \left(K^*/K^{*2}\right)^3$ est un homomorphisme.*

3.7. Lemme. *Le noyau de l'application ψ est égal à $2E(K)$.*

3.8. Lemme. *L'image $\psi(E(K))$ dans $\left(K^*/K^{*2}\right)^3$ est finie.*

Démonstration. (du lemme V-3.6) Si, P, Q, R sont trois points de la courbe E, l'égalité $P + Q + R = 0_E$ équivaut à dire que P, Q, R sont alignés. Soit alors $y = \lambda x + \mu$ l'équation de la droite D telle que l'intersection de D et E soit P, Q, R. Supposons d'abord que $\{P, Q, R\} \cap \{0_E, P_1, P_2, P_3\} = \emptyset$. L'équation $f(x) - (\lambda x + \mu)^2 = 0$ a donc pour racine $x(P)$, $x(Q)$ et $x(R)$.

Si l'on pose $x' = x - \alpha_i$ on trouve que

$$f(x' + \alpha_i) - (\lambda x' + \lambda\alpha_i + \mu)^2 = 0$$

a pour solutions $x(P)-\alpha_i$, $x(Q)-\alpha_i$ et $x(R)-\alpha_i$ et comme terme constant $-(\lambda\alpha_i + \mu)^2$ donc

$$(x(P) - \alpha_i)(x(Q) - \alpha_i)(x(R) - \alpha_i) = (\lambda\alpha_i + \mu)^2$$

et ainsi $\psi_i(P)\psi_i(Q)\psi_i(R) = 1$. On a bien ainsi que $R = P + Q$ entraîne $P, Q, -R$ alignés donc $\psi_i(P)\psi_i(Q)\psi_i(-R) = 1$; comme $\psi_i(-R) = \psi_i(R) = \psi_i(R)^{-1}$ on obtient bien $\psi_i(R) = \psi_i(P)\psi_i(Q)$. Cela achève la preuve si $\{P, Q, R\} \cap \{0_E, P_1, P_2, P_3\} = \emptyset$. Si $R = O_E$ la relation devient évidente. Sinon observons que $(x(P)-\alpha_1)(x(P)-\alpha_2)(x(P)-\alpha_3) = y(P)^2$; on vérifie cas par cas que la relation $\psi_i(P + Q) = \psi_i(P)\psi_i(Q)$ persiste toujours. □

Démonstration. (du lemme V-3.7) Il est clair que $2E(K) \subset \operatorname{Ker}\psi$ puisque K^*/K^{*2} est d'exposant 2. Il nous faut montrer que $\cap_i \operatorname{Ker}\psi_i \subset 2E(K)$. Supposons donc

$$\text{pour } i = 1, 2, 3,\ \exists z_i \in K^*,\ x(P) - \alpha_i = z_i^2. \tag{I.26}$$

Résolvons le système linéaire de type Vandermonde $u + v\alpha_i + w\alpha_i^2 = z_i$. On tire de l'égalité $(u + v\alpha_i + w\alpha_i^2)^2 = x - \alpha_i$ le système

$$\begin{cases} u^2 - 2vwb - x = 0 \\ 2uv - 2vwa - bw^2 + 1 = 0 \\ v^2 + 2uw - aw^2 = 0 \end{cases}$$

d'où l'on tire en particulier $v^3 + vw^2a + bw^3 - w = 0$ ou encore

$$\left(\frac{v}{w}\right)^3 + a\left(\frac{v}{w}\right) + b = \left(\frac{1}{w}\right)^2.$$

Ainsi $Q := \left(\frac{v}{w}, \frac{1}{w}\right) \in E(K)$. Un calcul direct montre alors que $P = 2Q$; en effet :

$$\begin{aligned} x(2Q) &= \frac{\left(\frac{v}{w}\right)^4 - 2a\left(\frac{v}{w}\right)^2 - 8b\left(\frac{v}{w}\right) + a^2}{4\left(\left(\frac{v}{w}\right)^3 + a\left(\frac{v}{w}\right) + b\right)} \\ &= \frac{v^4 - 2av^2w^2 - 8bvw^3 + a^2w^4}{4w^2} \\ &= \frac{aw^2 - 2uw)^2}{4w^2} + \frac{1}{4}(-2av^2 - 8bvw + aw^2) \\ &= u^2 - 2vwb + \frac{1}{4}(-2av^2 + 2a^2w^2 - 4auw) \\ &= x. \end{aligned}$$

□

Démonstration. (du lemme V-3.8) Choisissons un ensemble S fini de places

de K telles que ;
i) L'élément $2\Delta_E = 2(4a^3 + 27b^2)$ est une S-unité.
ii) L'anneau $\mathscr{O}_{K,S}$ est principal.

Notons que cela est possible d'après la remarque **??**. On peut maintenant écrire $x = A/B$, $y = C/D$ avec $A, B, C, D \in \mathscr{O}_{K,S}$ et $\mathrm{pgcd}(A, B) = \mathrm{pgcd}(C, D) = 1$ (dans l'anneau $\mathscr{O}_{K,S}$) ; l'équation $y^2 = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$ se traduit par $C^2B^3 = D^2(A-\alpha_1 B)(A-\alpha_2 B)(A-\alpha_3 B)$. Comme D premier avec C on a D^2 divise B^3 et comme B premier avec A, on a B^3 divise D^2, donc, quitte à modifier B et D par une unité on peut supposer $B^3 = D^2$ et disons $B = E^2$, $D = E^3$, ce qui donne

$$(x, y) = \left(\frac{A}{E^2}, \frac{C}{E^3}\right) \qquad \text{et} \qquad C^2 = (A-\alpha_1 E^2)(A-\alpha_2 E^2)(A-\alpha_3 E^2).$$

Si p (premier de $\mathscr{O}_{K,S}$) divise $(A-\alpha_1 E^2)$ et $(A-\alpha_2 E^2)$, il divise $(\alpha_1-\alpha_2)E^2$ et $(\alpha_1-\alpha_2)A$ donc $(\alpha_1-\alpha_2)$ qui est inversible. Les facteurs sont premiers entre eux donc sont des carrés modulo une unité. On obtient ainsi

$$x(P) - \alpha_i = \frac{A - \alpha_i E^2}{E^2} = \epsilon_i t_i^2.$$

où $\epsilon_i \in \mathscr{O}_{K,S}^*$. Comme corollaire du théorème des unités généralisé, on a que $\mathscr{O}_{K,S}^*/\mathscr{O}_{K,S}^{*2}$ est fini donc on peut choisir les ϵ_i dans un ensemble fini. Ainsi $\psi(P) = (\epsilon_1, \epsilon_2, \epsilon_3)$ prend un nombre fini de valeurs possibles dans $\left(K^*/K^{*2}\right)^3$. □

Terminons ce paragraphe en indiquant brièvement les modifications pour traiter une courbe $y^2 = f(x)$ sans quitter K le corps des coefficients du polynôme f. On introduit l'anneau $A := K[X]/(f(X))$ en notant α l'image de X ; on pose alors $\psi(P) = x(P) - \alpha$ à valeurs dans $G := A^*/A^{*2}$, si $x(P)$ n'est pas une racine de $f(X)$; on modifie comme précédemment la définition pour le cas particulier des points de 2-torsion. .

4. Le théorème de Siegel

On s'intéresse maintenant aux solutions entières ; le résultat principal que nous allons exposer est le suivant.

4.1. Théorème. *(Siegel) Soit C une courbe affine d'équation*

$$y^2 = f(x) = x^3 + ax + b$$

avec $a, b \in \mathscr{O}_K$ et $\Delta := 4a^3 + 27b^2 \neq 0$, alors l'ensemble des points $P = (x, y)$ sur la courbe avec $x, y \in \mathscr{O}_K$ est fini.

4.2. Remarque. La condition de lissité est indispensable car par exemple la courbe $y^2 = x^3$ possède toutes les solutions (t^2, t^3) (avec $t \in \mathscr{O}_K$) tandis que la courbe $y^2 = x^3 - x^2$ possède toutes les solutions (t^2+1, t^3+t) (avec $t \in \mathscr{O}_K$). On peut déduire du théorème précédent (mais nous ne le ferons pas) un théorème apparemment plus général :

4.3. Théorème. *(Siegel) Soit C une courbe affine d'équation*

$$ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0$$

telle que la courbe projective correspondante soit lisse, alors l'ensemble des points $P = (x, y)$ sur la courbe avec x, y entiers (algébriques) est fini.

Nous allons déduire le théorème de Siegel du résultat suivant, également dû à Siegel.

4.4. Théorème. *(Équation aux S-unités) Soit K un corps de nombres et S un ensemble fini de places. L'ensemble des couples de S-unités $(x, y) \in (\mathscr{O}_{K,S}^*)^2$ vérifiant*

$$x + y = 1 \tag{I.27}$$

est un ensemble fini.

Démonstration. (Réduction du théorème V-4.1 au théorème V-4.4) On peut si l'on veut augmenter l'ensemble S et le corps K. On peut donc supposer $\mathscr{O}_{K,S}$ principal, $\Delta \in \mathscr{O}_{K,S}^*$ et $f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$. Soit donc $(x, y) \in (\mathscr{O}_{K,S})^2$ solution entière. Comme dans la démonstration du théorème de Mordell-Weil, on en déduit une factorisation :

$$x - \alpha_i = b_i z_i^2$$

avec b_i représentants de $\mathscr{O}_{K,S}^*/\mathscr{O}_{K,S}^{*2}$. Introduisons les nombres algébriques $\beta_i = \sqrt{b_i}$ qui sont dans K' extension finie de K. On tire alors de $x - \alpha_i = (\beta_i z_i)^2$ les relations

$$\alpha_i - \alpha_j = (\beta_i z_i - \beta_j)(\beta_i z_i + \beta_j z_j) \in \mathscr{O}_{K,S}^*.$$

On en déduit les « identités de Siegel » :

$$\frac{\beta_i z_i \pm \beta_j z_j}{\beta_i z_i - \beta_k z_k} \mp \frac{\beta_j z_j \pm \beta_k z_k}{\beta_i z_i - \beta_k z_k} = 1. \tag{I.28}$$

D'après le théorème V-4.4 l'ensemble des valeurs prises par $\epsilon := \frac{\beta_i z_i \pm \beta_j z_j}{\beta_i z_i - \beta_k z_k}$ est fini, et on en tire aisément que les valeurs $\beta_i z_i$ sont en nombre fini et donc de même les valeurs de x et donc de y. □

4.5. Remarque. On peut observer que la réduction est effective au sens suivant : si l'on dispose d'un algorithme pour calculer les solutions de l'équation aux S-unités, on disposera d'un algorithme pour calculer l'ensemble des solutions entiéres de $y^2 = f(x)$.

On connait essentiellement deux preuves du théorème V-4.4, celle de Siegel, basée sur un théorème d'approximation rationnelle, et celle de Baker basée sur son théorème sur les formes linéaires de logarithmes. Le défaut de la preuve de Siegel est d'être ineffective : elle ne permet pas la détermination de l'ensemble fini de solutions ; c'est celle que nous exposerons néanmoins, pour une esquisse de la théorie de Baker, voir le chapitre suivant, section 4.

Réduction. Soit $m \geqslant 2$, on sait, d'après le théorème des unités généralisé, que le groupe $\mathscr{O}_{K,S}^*/\mathscr{O}_{K,S}^{*m}$ est fini. En d'autres termes il existe un ensemble fini de S-unités ϵ_i tel que toute S-unité s'écrive $x = \epsilon_i z^m$ avec $z \in \mathscr{O}_{K,S}^*$. Ainsi les solutions (x, y) de l'équation aux S-unités fournissent des solutions d'une des équations (en nombre fini) suivantes :

$$\epsilon_1 z_1^m + \epsilon_2 z_2^m = 1$$

et il suffit de prouver que ces dernières n'ont qu'un nombre fini de solutions $(z_1, z_2) \in (\mathscr{O}_{K,S}^*)^2$ ou même dans $(\mathscr{O}_{K,S})^2$.

4.6. Proposition. *Soient $a, b \in \mathscr{O}_{K,S}$ et $m \geqslant 3$. L'ensemble des solutions S-entières de l'équation*

$$ax^m + by^m = 1$$

est fini.

Démonstration. Nous donnons la preuve dans le cas le plus simple pour les notations, le cas $\mathscr{O}_{K,S} = \mathbf{Z}$. Si l'on note $\alpha = \sqrt[m]{-\frac{b}{a}}$ on obtient

$$\left(\frac{x}{y}\right)^m + \frac{b}{a} = \left(\frac{x}{y} - \alpha\right) F\left(\frac{x}{y}\right) = \frac{1}{ay^m}$$

En observant que si x/y est proche de α, il doit être à une distance bornée inférieurement des autres racines (celles de F), on en tire donc une inégalité du type

$$\left|\frac{x}{y} - \alpha\right| \leqslant \frac{C_1}{|y|^m}, \tag{I.29}$$

où la constante C_1 ne dépend que de α. Pour conclure il suffit de disposer d'un énoncé d'approximation diophantienne du type

$$\frac{C_2}{|y|^\delta} \leqslant \left|\frac{x}{y} - \alpha\right|, \tag{I.30}$$

avec C_2 dépendant de α et δ et surtout $\delta < m$. En effet en combinant les inégalités (V-V.29) et (V-V.30) on obtient

$$|y| \leqslant \left(\frac{C_1}{C_2}\right)^{\frac{1}{m-\delta}}.$$

Un énoncé du type de l'inégalité (V-V.30) est fourni par le théorème de Roth qui permet de choisir n'importe quel $\delta > 2$. Toutefois la preuve du théorème de Roth est ineffective au sens où elle ne permet pas de calculer la constante $C_2 = C_2(\alpha, \delta)$. Une méthode donnant des résultats effectifs a été développée dans les années soixante par Baker et sera brièvement discutée au chapitre suivant. □

5. Courbes elliptiques sur les complexes

On décrit dans ce paragraphe le lien avec la théorie classique des fonctions elliptiques, justifiant ainsi le nom de « courbes elliptiques », les fonctions elliptiques tirant leur nom du fait qu'elle interviennent dans le calcul de la longueur d'un arc d'ellipse.

Nous aurons besoin du résultat classique de variable complexe suivant

5.1. Théorème. *(Liouville) Un fonction entière (i.e. holomorphe sur* $\mathbf{C}$ *tout entier) et bornée est constante.*

Considérons maintenant $\Omega := \mathbf{Z}\omega_1 \oplus \mathbf{Z}\omega_2$ un réseau dans $\mathbf{C}$ et étudions les fonctions Ω-périodique, i.e. telles que $f(z+\omega) = f(z)$ pour $\omega \in \Omega$. Le théorème de Liouville indique tout de suite que les seules fonctions entières et Ω-périodiques sont les constantes, ce qui justifie la définition suivante.

5.2. Définition. Une fonction *elliptique* est une fonction méromorphe sur $\mathbf{C}$ et Ω-périodique pour un réseau Ω.

Remarquons que l'ensemble des fonctions Ω-elliptiques forme un corps qu'on notera $\mathscr{M}(\Omega)$, qui contient les constantes, i.e. le corps $\mathbf{C}$, et est stable par dérivation. Voyons tout de suite que cette définition n'est pas réduite aux constantes.

5.3. Définition. Soit $\Omega := \mathbf{Z}\omega_1 \oplus \mathbf{Z}\omega_2$ un réseau dans $\mathbf{C}$, on définit la *fonction de Weierstrass* associée à Ω par la formule :

$$\wp(z) = \wp(z;\Omega) = \frac{1}{z^2} + \sum_{\omega\in\Omega}{}' \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right), \tag{I.31}$$

où le signe $\sum'$ signifie qu'on omet $\omega = 0$.

La fonction de Weierstrass permet de donner une description complète des fonctions elliptiques et d'établir le lien avec les courbes elliptiques.

5.4. Théorème. *La fonction $\wp$ de Weierstrass est une fonction elliptique. Le corps des fonctions Ω-elliptiques est engendré par $\wp$ et sa dérivée $\wp'$, i.e. $\mathscr{M}(\Omega) = \mathbf{C}(\wp, \wp')$. De plus ces deux fonctions elliptiques sont liées par la relation algébrique :*

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3 \tag{I.32}$$

où les constantes g_2 et g_3 sont définies par

$$g_2 = g_2(\Omega) = 60\sum_{\omega\in\Omega}{}' \frac{1}{\omega^4} \qquad \text{et} \qquad g_3 = g_3(\Omega) = 140\sum_{\omega\in\Omega}{}' \frac{1}{\omega^6}$$

Démonstration. (esquisse) La série définissant la dérivée

$$\wp'(z) = -2\sum_{\omega\in\Omega} \frac{1}{(z-\omega)^3}$$

est absolument convergente, uniformément sur tout compact évitant Ω ; elle définit donc une fonction holomorphe sur $\mathbf{C}\setminus\Omega$ qui est visiblement Ω-périodique et impaire. De plus $\wp'$ a un pôle d'ordre 3 en chaque point de Ω, on a donc bien $\wp' \in \mathscr{M}(\Omega)$. La série définissant $\wp$ montre qu'elle est méromorphe avec un pôle double en chaque $\omega \in \Omega$ et est paire. De la périodicité de $\wp'$ on tire que $\wp(z+\omega) = \wp(z) + C_\omega$. Soit ω un des générateurs de Ω de sorte que $\omega/2 \notin \Omega$, en prenant $z := -\omega/2$ on obtient $\wp(-\omega/2) = \wp(\omega/2) = \wp(-\omega/2) + C_\omega$ d'où $C_\omega = 0$, ainsi on a aussi $\wp \in \mathscr{M}(\Omega)$. Pour montrer que $\mathscr{M}(\Omega) = \mathbf{C}(\wp, \wp')$, on décompose une fonction de $\mathscr{M}(\Omega)$ en paire + impaire, et on se ramène à prouver qu'une fonction f qui est Ω-elliptique et paire est dans $\mathbf{C}(\wp)$. Pour cela on prouve que ses pôles et zéros sont tous d'ordres pairs et donc f a les mêmes zéros et pôles qu'une fraction du type $\prod_i (\wp(z) - \wp(u_i))^{m_i}$ et donc que les deux fonctions coïncident à une constante près. Pour prouver la relation de dépendance algébrique, on calcule le développement de Taylor de $\wp(z)$ (ou plutôt de

$\wp(z) - z^{-2})$ en $z = 0$:

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} a_n z^n \qquad \text{avec} \quad a_n = (n+1) \sum_{\omega \in \Omega}{}' \frac{1}{\omega^{n+2}}. \tag{I.33}$$

Un calcul du développement de Taylor (uniquement les pôles et le terme constant) de la fonction $\psi(z) = \wp'(z)^2 - 4\wp(z)^3 + g_2\wp(z) + g_3$ montre qu'elle est holomorphe et nulle en 0 ; par le théorème de Liouville la fonction $\psi(z)$ est donc identiquement nulle. □

5.5. Corollaire. *Soit Ω un réseau de $\mathbf{C}$ l'application $z \mapsto (1, \wp(z), \wp'(z))$ étendue par $\omega \mapsto (0, 0, 1)$ définit une application biholomorphe de $\mathbf{C}/\Omega$ vers la cubique projective d'équation (dans $\mathbf{P}^2$) :*

$$TY^2 = 4X^3 - g_2XT^2 - g_3T^3.$$

De plus l'application est un isomorphisme de groupes.

Démonstration. La première affirmation résulte essentiellement du théorème précédent. La deuxième affirmation découle de la comparaison de la loi d'addition algébrique et de la formule d'addition suivante concernant la fonction de Weierstrass :

$$\wp(u+v) = -\wp(u) - \wp(v) + \frac{1}{4}\left(\frac{\wp'(u) - \wp'(v)}{\wp(u) - \wp(v)}\right)^2. \tag{I.34}$$

Cette dernière formule peut se démontrer en vérifiant que, pour v fixé, les pôles du membre de gauche sont des pôles doubles en chaque $u \in -v + \Omega$; le membre de droite a en fait les mêmes pôles car $\frac{\wp'(u)-\wp'(v)}{\wp(u)-\wp(v)}$ a un pôle simple pour $u \in \Omega$ qui va se compenser avec le terme $-\wp(u)$ et comme $\wp(u) - \wp(v) = 0$ si et seulement si $u \pm v \in \Omega$ mais $\wp'(u) - \wp'(v)$ s'annule pour $u \in v + \Omega$. Une fois vérifiée l'égalité des termes avec pôles, la formule V-V.34 suit. □

Inversement, on peut montrer que, étant donnés $g_2, g_3 \in \mathbf{C}$ vérifiant $\Delta := g_2^3 - 27g_3^2 \neq 0$, il existe un réseau Ω tel que $g_2 = g_2(\Omega)$ et $g_3 = g_3(\omega)$. Ainsi, sur le corps des complexes, on peut voir une courbe elliptique comme un tore complexe, i.e. un quotient $\mathbf{C}/\Omega$. Ce point de vue permet de voir clairement plusieurs propriétés des courbes elliptiques ou des familles de courbes elliptiques ; les deux propositions suivantes donnent des illustrations de ce principe.

5.6. Proposition. *Soit $E = \mathbf{C}/\Omega$ une courbe elliptique, on a alors*

$$\mathrm{Ker}[N]_E = \frac{1}{N}\Omega/\Omega \cong (\mathbf{Z}/N\mathbf{Z})^2 \tag{I.35}$$

Démonstration. En effet l'application $[N]_E : \mathbf{C}/\Omega \to \mathbf{C}/\Omega$ est induite par la multiplication par N sur $\mathbf{C}$ donc $\mathrm{Ker}[N]_E = \{z \in \mathbf{C} \mid Nz \in \Omega\}/\Omega$ et comme $\Omega \cong \mathbf{Z}^2$ le résultat est clair. □

5.7. Remarque. On peut observer que les points de torsion permettent de reconstruire en partie le réseau Ω ; plus précisément on voit aisément que, pour tout nombre premier ℓ, on a :

$$\lim_n \mathrm{Ker}[\ell^n] = \lim_n \Omega/\ell^n\Omega \cong \left(\lim_n \mathbf{Z}/\ell^n\mathbf{Z}\right)^2.$$

Ainsi, en introduisant l'anneau $\mathbf{Z}_\ell := \lim_n \mathbf{Z}/\ell^n\mathbf{Z}$ on voit que

$$\lim_n \mathrm{Ker}[\ell^n] = \Omega \otimes \mathbf{Z}_\ell.$$

Cette remarque peut paraître pédante quand on s'occupe de courbes elliptiques sur $\mathbf{C}$ mais elle devient fondamentale sur un corps quelconque (par exemple un corps fini) car le membre de gauche garde un sens, alors que le membre de droite (disons Ω) n'existe pas. Plus précisément on peut définir :

5.8. Définition. Soit E une courbe elliptique définie sur un corps K et ℓ un nombre premier ne distinct de la caractéristique de K. Le *module de Tate* ℓ-adique d'une courbe elliptique est défini par

$$T_\ell(E) := \lim_{\leftarrow} E[\ell^n].$$

5.9. Remarque. Notons que, si $u \in \mathbf{C}^*$, alors $\mathbf{C}/\Omega$ est isomorphe $\mathbf{C}/u\Omega$ (par la multiplication par u). Par ailleurs on vérifie aisément que $u^2\wp(uz, u\Omega) = \wp(z, \Omega)$ et $u^3\wp'(uz, u\Omega) = \wp'(z, \Omega)$. On peut donc toujours, à isomomorphisme près, remplacer le réseau $\mathbf{Z}\omega_1 \oplus \mathbf{Z}\omega_2$ par le réseau homothétique $\mathbf{Z} \oplus \mathbf{Z}\tau$ où l'on a posé $\tau := \frac{\omega_2}{\omega_1}$. Quitte à échanger ω_1 et ω_2, on peut de plus supposer que $\mathrm{Im}(\tau) > 0$. Toute courbe elliptique (sur $\mathbf{C}$) est donc isomorphe à un tore $\mathbf{C}/\mathbf{Z} \oplus \mathbf{Z}\tau$ avec τ dans le *demi-plan de Poincaré*

$$\mathscr{H} := \{\tau \in \mathbf{C} \mid \mathrm{Im}(\tau) > 0\}.$$

Le résultat suivant précise quand deux telles courbes sont isomorphes.

5.10. Proposition. *Deux tores $E_\tau = \mathbf{C}/\mathbf{Z} \oplus \mathbf{Z}\tau$ et $E_{\tau'} = \mathbf{C}/\mathbf{Z} \oplus \mathbf{Z}\tau'$ sont isomorphes si et seulement si il existe $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbf{Z})$ telles que*

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

En particulier on peut identifier l'espace des classes d'isomorphisme de courbes elliptiques complexes avec l'espace $\mathrm{SL}(2,\mathbf{Z})\backslash\mathscr{H}$.

Démonstration. Un homomorphisme $\phi : \mathbf{C}/\mathbf{Z}\oplus\mathbf{Z}\tau \to \mathbf{C}/\mathbf{Z}\oplus\mathbf{Z}\tau'$ est induit par un homomorphisme de $\mathbf{C}$ vers $\mathbf{C}$, i.e. par la multiplication par $\alpha \in \mathbf{C}$ tel que $\alpha(\mathbf{Z}\oplus\mathbf{Z}\tau) \subset \mathbf{Z}\oplus\mathbf{Z}\tau'$. En particulier $\alpha = c\tau' + d$ (avec $c, d \in \mathbf{Z}$) et $\alpha\tau = a\tau' + b$ (avec $a, b \in \mathbf{Z}$). Ainsi on en tire bien $\tau = (a\tau' + b)/(c\tau' + d)$. Le fait que ϕ soit un isomorphisme se traduit par $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2,\mathbf{Z})$ et comme $\mathrm{Im}(\tau) = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mathrm{Im}(\tau')/|c\tau' + d|^2$ on voit qu'en fait la matrice est dans $\mathrm{SL}(2,\mathbf{Z})$. □

5.11. Proposition. *Soit* $E = \mathbf{C}/\Omega$ *une courbe elliptique ; supposons* $\Omega = \mathbf{Z} + \mathbf{Z}\tau$ *on a alors*

$$\mathrm{End}(E) = \{\alpha \in \mathbf{C} \mid \alpha\Omega \subset \Omega\} = \begin{cases} \mathbf{Z} & \text{si } [\mathbf{Q}(\tau) : \mathbf{Q}] > 2 \\ \mathbf{Z} + \mathbf{Z}A\tau & \text{si } [\mathbf{Q}(\tau) : \mathbf{Q}] = 2 \end{cases} \tag{I.36}$$

où, dans le second cas, l'entier A *est le coefficient dominant de l'équation minimale* $A\tau^2 + B\tau + C = 0$.

Dans le deuxième cas où $\mathrm{End}(E)$ est un sous-anneau d'indice fini dans l'anneau des entiers d'un corps quadratique imaginaire on dit que E admet des « *multiplications complexes* », ou encore en abrégé est de « *type CM* ».

Démonstration. En reprenant le calcul de la démonstration précédente, un endomorphisme est donné par la multiplication par $\alpha = c\tau + d$ et correspond à une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbf{Z})$ telle que $\tau = (a\tau + b)/(c\tau + d)$ ou encore $c\tau^2 + (d - a)\tau - b = 0$. Si $[\mathbf{Q}(\tau) : \mathbf{Q}] > 2$ la seule possibilité est $c = b = 0$ et $a = d$, c'est-à-dire la multiplication par d ; si τ est quadratique et vérifie l'équation minimale $A\tau^2 + B\tau + C = 0$, on en tire $c = mA$, $d - a = mB$ et $-b = mC$ d'où $\alpha = mA\tau + d \in \mathbf{Z} + \mathbf{Z}A\tau$. □

5.12. Remarque. Remarquons que, si l'on voit $\mathrm{End}(E)$ comme un sous-anneau de $\mathbf{C}$ on peut facilement vérifier la formule suivante

$$\deg(\alpha) := \mathrm{card}\,\mathrm{Ker}(\alpha) = \mathrm{N}(\alpha) = \alpha\bar{\alpha}. \tag{I.37}$$

En particulier l'application $\deg : \mathrm{End}(E) :\to \mathbf{Z}$ est quadratique.

$$T_\ell(E) := \lim_{\leftarrow} E[\ell^n]$$

6. Courbes elliptiques sur un corps fini

On va, dans ce paragraphe, transposer un certain nombre de résultats du paragraphe précédent aux corps de caractéristique p et notamment aux corps finis. Je renvoie au livre de Silverman [?] pour les preuves complètes. Les courbes elliptiques sur les corps finis sont notamment utilisées en cryptographie, voir par exemple l'ouvrage [?].

Soit E une courbe projective plane d'équation V-V.2 avec $a_i \in \mathbf{F}_q$, le groupe des points rationnels $E(\mathbf{F}_q)$ est évidemment fini et, en particulier, tous les points sont des points de torsion. Une telle courbe possède encore les applications « multiplication par un entier n » mais elle possède aussi un endomorphisme remarquable spécifique à la caractéristique p.

6.1. Définition. L'endomorphisme « Frobenius » de $E/\mathbf{F}_q$ est défini par la formule :

$$\Phi_q(x, y) = (x^q, y^q).$$

Remarquons que si $f(x, y) = 0$ (avec $f(X, Y) \in \mathbf{F}_q[X, Y]$) alors $f(x^q, y^q) = (f(x, y))^q = 0$ donc Φ_q est bien un endomorphisme (il est clair qu'il va aussi respecter la loi d'addition).

Nous admettrons (voir par exemple [?]) l'énoncé suivant, analogue de la proposition V-5.6 :

6.2. Proposition. *Soit E une courbe elliptique elliptique définie sur un corps fini$\mathbf{F}_q$ et N un entier $\geqslant 2$. Notons* $\mathrm{Ker}[N]_E := \{P \in E(\bar{\mathbf{F}}_q) \mid NP = O_E\}$*, on a alors, si N n'est pas divisible par la caractéristique du corps :*

$$\mathrm{Ker}[N]_E = \frac{1}{N}\Omega/\Omega \cong (\mathbf{Z}/N\mathbf{Z})^2 \tag{I.38}$$

6.3. Remarques. En particulier, pour ℓ premier distinct de la caractéristique de $\mathbf{F}_q$, on peut définir le module de Tate comme sur le corps des complexes

$$T_\ell(E) := \varprojlim \mathrm{Ker}[\ell^n]_E.$$

On a de nouveau $T_\ell(E) \cong (\mathbf{Z}_\ell)^2$ (comme $\mathbf{Z}_\ell$-module) ; noter qu'on ne dispose pas, par contre d'un réseau naturel $\Omega \cong \mathbf{Z}^2$ tel que $T_\ell(E) \cong \Omega \otimes \mathbf{Z}_\ell$.

L'énoncé de la proposition n'est plus vrai pour $N = p^m$ avec p la caractéristique de $\mathbf{F}_q$. On peut montrer qu'on a dans ce cas, ou bien $\mathrm{Ker}[p^m]_E \cong \mathbf{Z}/p^m\mathbf{Z}$ (cas « *ordinaire* ») ou bien $\mathrm{Ker}[p^m]_E = \{O_E\}$ (cas « *supersingulier* »).

Le résultat clef concernant le nombre des points rationnels sur $\mathbf{F}_q$ est le suivant.

6.4. Théorème. *(Hasse) Soit E une courbe elliptique définie sur $\mathbf{F}_q$ alors*

$$|\operatorname{card} E(\mathbf{F}_q) - q - 1| \leqslant 2\sqrt{q} \tag{I.39}$$

Plus précisément, il existe α entier algébrique vérifiant $\alpha\bar{\alpha} = q$ tel que

$$\operatorname{card} E(\mathbf{F}_{q^m}) = q^m + 1 - \alpha^m - \bar{\alpha}^m \tag{I.40}$$

Démonstration. (partiellement admis) L'ensemble des points rationnels est aussi l'ensemble des points fixes du Frobenius Φ_q. Admettons que, comme sur les complexes, le degré d'un endomorphisme soit donné par une fonction quadratique et donc qu'en particulier :

$$\deg(n\Phi_q - m) = P(n,m) = an^2 + 2bmn + cm^2.$$

On a alors d'une part $\operatorname{card} E(\mathbf{F}_q) = P(1,1) = a + c - 2b$ et d'autre part $c = P(0,1) = 1$ et $a = P(1,0) = q$. Enfin, comme le polynôme $P(n,m)$ est à valeurs positives, on en tire que $b^2 - ac \leqslant 0$ et donc $|b| \leqslant \sqrt{q}$, ce qui démontre la première formule.

Pour la deuxième on peut donner une deuxième interprétation de la formule précédente en invoquant une analogie avec la situation sur les complexes. Alors on peut interpréter Φ_q comme un endomorphisme du module de Tate dont les valeurs propres α, $\bar{\alpha}$ vérifient $\alpha\bar{\alpha} = q$; les valeurs propres de Φ_q^m sont alors α^m et $\bar{\alpha}^m$, d'où la deuxième formule. □

7. La fonction L d'une courbe elliptique

Soit E une courbe elliptique sur $\mathbf{Q}$, on peut supposer qu'elle est donnée comme projective plane d'équation V-V.2, avec $a_i \in \mathbf{Z}$ et on peut supposer cette équation *minimale*, c'est-à-dire que si Δ_E est son discriminant, le discriminant de tout autre équation à coefficients entiers sera de la forme $\Delta = u^{12}\Delta_E$, avec $u \in \mathbf{Z}$.

Pour simplifier nous continuons la discussion en restant sur le corps des rationnels. La réduction modulo p possède au plus un point singulier qui est de type point de rebroussement $y^2 = x^3$ ou point de croisement $y^2 + axy + bx^2 = x^3$, le polynôme $y^2 + axy + bx^2 = (y - \alpha x)(y - \alpha' x)$ étant irréductible (si $\alpha \in \mathbf{F}_{p^2} \setminus \mathbf{F}_p$) ou non (si $\alpha \in \mathbf{F}_p$) sur $\mathbf{F}_p$. Géométriquement, dans le deuxième cas, le cône tangent au point singulier est composé de deux droites $y = \alpha x$ et $y = \alpha' x$ qui peuvent être rationnelles sur $\mathbf{F}_p$ ou définies sur $\mathbf{F}_{p^2}$ et conjuguées.

7.1. Définition. Soit p premier et $E/\mathbf{Q}$ ayant pour modèle minimal

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

On dit que E a *bonne réduction* en p si ce modèle reste lisse modulo p (i.e. si p ne divise pas Δ_E). On dit que E a *réduction additive* en p si ce modèle est singulier modulo p et la singularité a une tangente unique (i.e. si p divise c_4 et Δ_E) et *réduction multiplicative* en p si ce modèle est singulier modulo p avec deux tangentes distinctes (i.e. si p ne divise pas c_4 mais divise Δ_E) ; si les deux tangentes sont définies sur $\mathbf{F}_p$ (resp. ne sont pas définies sur $\mathbf{F}_p$) on dit que la réduction est multiplicative *déployée* (resp. *non déployée*).

7.2. Remarque. L'appellation « additive » ou « multiplicative » provient de l'observation suivante dont la vérification est laissée au lecteur. Si E une cubique de Weierstrass est singulière (nécessairement en un point unique P_0), alors le procédé de cordes et tangentes définit encore une loi de groupe sur $E \setminus \{P_0\}$ et ce groupe est isomorphe au groupe additif si le point singulier est un point de rebroussement et au groupe multiplicatif si les tangentes sont distinctes. Plus précisément $E(K) \setminus \{P_0\}$ est isomorphe à $(K, +)$ si la réduction est additive, à $(K^*, \times)$ si la rédution est multiplicative déployée et à $(K_1, \times)$ si la réduction est multiplicative non déployée où $K_1 = \mathrm{Ker}\left\{\mathrm{N}_K^{K'} : K'^* \to K^*\right\}$ avec K' extension quadratique.

7.3. Définition. Le *conducteur* de $E/\mathbf{Q}$ est défini comme $N_E := \prod_p p^{n(E,p)}$ avec

$$n(E,p) = \begin{cases} 0 & \text{si } E \text{ a bonne réduction en } p, \\ 1 & \text{si } E \text{ a réduction multiplicative en } p, \\ 2 + \delta_{E,p} & \text{si } E \text{ a réduction additive en } p, \end{cases}$$

avec $\delta_{E,p} = 0$ si $p \geqslant 5$ et $\delta_{E,2} \leqslant 8$, $\delta_{E,3} \leqslant 5$

7.4. Définition. On définit la fonction $L(E,s)$ et ses facteurs locaux par :

$$L_p(E,s) = \begin{cases} \left(1 - a_p p^{-s} + p^{1-2s}\right)^{-1} & \text{si } p \text{ ne divise pas } \Delta_E \\ (1-p^{-1})^{-1} & \text{si } E \text{ a réduction multiplicative déployée} \\ (1+p^{-1})^{-1} & \text{si } E \text{ a réduction multiplicative non déployée} \\ 1 & \text{si } E \text{ a réduction additive} \end{cases} \tag{I.41}$$

$$L(E,s) = \sum_{n=1}^{\infty} a_n n^{-s} = \prod_p L_p(E,s) \tag{I.42}$$

7.5. Proposition. *La série de Dirichlet et le produit d'Euler définissant*

la fonction $L(E,s)$ sont absolument convergents pour $\mathrm{Re}(s) > 3/2$.

Démonstration. C'est immédiat à partir du théorème de Hasse. □

7.6. Théorème. *(Wiles [?]) La fonction $L(E,s)$ se prolonge analytiquement en une fonction entière qui vérifie l'équation fonctionnelle suivante, où l'on pose $\Lambda(E,s) := N_E^{s/2}(2\pi)^{-s}\Gamma(s)L(E,s)$:*

$$\Lambda(E,s) = \pm\Lambda(E,2-s) \tag{I.43}$$

On observera bien sûr l'analogie avec l'équation fonctionnelle de la fonction zêta de Riemann. Le théorème de Wiles est en fait plus précis et « explique » l'équation fonctionnelle de la série de Dirichlet $L(E,s) = \sum_{n=1}^{\infty} a_n n^{-s}$ par le fait que la fonction associée (pour z dans le demi-plan de Poincaré) $f_E(z) = \sum_{n=1}^{\infty} a_n \exp(2\pi i n z)$ est *modulaire* de niveau N_E et poids 2, et vérifie l'équation fonctionnelle suivante :

$$f_E\left(-\frac{1}{N_E z}\right) = \pm N_E z^2 f_E(z). \tag{I.44}$$

Un calcul formel analogue à celui effectué pour démontrer l'équation fonctionnelle de la fonction $\zeta(s)$ montre que l'équation V-V.44 pour f_E entraîne l'équation fonctionelle pour $L(E,s)$.

Le lien entre le théorème de Wiles et le « théorème » de Fermat est le suivant. En partant d'une hypothétique solution de l'équation de Fermat

$$a^\ell + b^\ell + c^\ell = 0,$$

on fabrique une courbe elliptique, dite courbe de Frey :

$$y^2 = x(x-a^\ell)(x+b^\ell).$$

En examinant cette hypothétique courbe, on s'aperçoit qu'elle a (aurait) des propriétés remarquables. Par exemple $\Delta_E = 2^{-8}(abc)^{2\ell}$ et $N_E = \prod_{p\,|\,abc} p$. La forme modulaire associée (par le théorème de Wiles) aurait des propriétés encore plus extraordinaires : par un théorème de Ribet [?] on pourrait abaisser son niveau de N_E jusqu'à 2. Or il n'existe pas de telle forme non nulle de niveau 2, ce qui achève la démonstration du théorème de Fermat !

Pour énoncer la conjecture suivante, nous aurons besoin de définir la *période* réelle de E

$$\Omega_E := \int_{E(\mathbf{R})} \frac{dx}{2y + a_1 x + a_3} \tag{I.45}$$

7.7. Conjecture. *(Birch & Swinnerton-Dyer)*

I) L'ordre d'annulation de la fonction $L(E,s)$ en $s=1$ est égal au rang du groupe de Mordell-Weil $r = \mathrm{rang}\, E(\mathbf{Q})$.

II) Soit $P_1, \ldots, P_r$ une base de $E(\mathbf{Q})$ modulo torsion, soit Ω_E la période de E, le coefficient dominant de $L(E,s)$ en $s=1$ est donné par

$$\lim_{s\to 1} \frac{L(E,s)}{(s-1)^r} = u\Omega_E \det\left(\langle P_i, P_j\rangle\right) \tag{I.46}$$

où $u \in \mathbf{Q}^*$.

7.8. Remarques. Dans la formulation complète[2] le nombre u a une définition précise. En fait

$$u = \frac{M \prod_{p \,|\, \Delta_E} c_p}{|E(\mathbf{Q})_{\mathrm{tor}}|^2}$$

où $c_p \geqslant 1$ est un entier dépendant de la mauvaise réduction en p et M est le cardinal du « groupe de Tate-Shafarevic » qui devrait être fini (c'est démontré seulement dans certains cas) et être un carré parfait (c'est vrai dès que le groupe en question est fini).

L'observation première de Birch & Swinnerton-Dyer est que, au moins formellement, $L(1) = \prod_p \frac{p}{N_p}$ où $N_p := \operatorname{card} E(\mathbf{F}_p)$. Or N_p vaut environ p avec une variation d'au plus $2\sqrt{p}$, notons cela $N_p = p + \delta(p)\sqrt{p}$, ainsi, toujours formellement, $L(1) = \prod_p (1+\delta(p)p^{-1/2})^{-1}$. Si $E(\mathbf{Q})$ est fini, on peut imaginer que N_p oscillera régulièrement et donc que $\delta(p)$ aura tendance à être bien réparti dans l'intervalle $[-2,2]$, ce qui ferait converger le produit ; si maintenant, $E(\mathbf{Q})$ est infini, on peut penser qu'on trouvera plus de points modulo p et donc que $\delta(p)$ aura tendance à être positif, ce qui entraînerait la divergence du produit et plus précisément forcerait $L(1)$ à être nulle.

On peut voir la conjecture comme une version sophistiquée du principe local/global. En effet la fonction $L(E,s)$ est construite à partir d'information locale assez simple, essentiellement le nombre de points modulo p, et grâce au prolongement analytique permettrait de retrouver le rang du groupe $E(\mathbf{Q})$.

Enfin le signe de l'équation fonctionnelle de $L(E,s)$ détermine la parité de l'ordre du zéro de $L(E,s)$ en $s=1$. Ainsi, conjecturalement le signe de l'équation fonctionnele détermine la parité du rang du groupe $E(\mathbf{Q})$. Cette version affaiblie s'appelle la *conjecture de parité*.

[2] La conjecture de Birch & Swinnerton-Dyer est un des problèmes pour la résolution desquels la fondation Clay propose un million de dollars.