Points entiers et rationnels sur des courbes et variétés modulaires de dimension supérieure

Samuel Le Fourn

université de BORDEAUX

THÈSE

présentée à

L'UNIVERSITÉ DE BORDEAUX

École Doctorale De Mathématiques et Informatique

par Samuel LE FOURN

sous la direction de Pierre PARENT

pour obtenir le grade de

DOCTEUR

Spécialité : Mathématiques Pures

⋆⋆⋆

Points entiers et rationnels sur des courbes et variétés modulaires de dimension supérieure

⋆⋆⋆

Soutenue le 20 novembre 2015 à l'Institut de Mathématiques de Bordeaux

après avis de

Aaron Levin	Assistant Professor, Michigan State University	Rapporteur
René Schoof	Professeur, Universita di Roma	Rapporteur

devant le jury de thèse composé de

Pascal Autissier	Professeur, Université de Bordeaux	Examinateur
Yuri Bilu	Professeur, Université de Bordeaux	Examinateur
Éric Gaudron	Professeur, Université Blaise Pascal	Examinateur
Loïc Merel	Professeur, Université Paris-Diderot	Examinateur
Pierre Parent	Maître de conférences, Université de Bordeaux	Directeur
René Schoof	Professeur, Universita di Roma	Rapporteur

Points entiers et rationnels sur des courbes et variétés modulaires de dimension supérieure

Samuel Le Fourn

2015

Institut de Mathématiques de Bordeaux
Université de Bordeaux
351, cours de la Libération, 33405 Talence cedex

Résumé

Cette thèse porte sur l'étude des points entiers et rationnels de certaines courbes et variétés modulaires. Après une brève introduction décrivant les motivations et le cadre de ce genre d'études ainsi que les résultats principaux de la thèse, le manuscrit se divise en trois parties.

Le premier chapitre s'intéresse aux $\mathbb{Q}$-courbes, et aux morphismes $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}_2(\mathbb{F}_p)$ qu'on peut leur associer pour tout p premier. Nous montrons que sous de bonnes hypothèses, pour p assez grand par rapport au discriminant du corps de définition de la $\mathbb{Q}$-courbe, ce morphisme est surjectif, ce qui résout un cas particulier du problème d'uniformité de Serre (toujours ouvert en général). Les outils principaux du chapitre sont la méthode de Mazur (basée ici sur des résultats d'Ellenberg), la méthode de Runge et des théorèmes d'isogénie, suivant la structure de preuve de Bilu et Parent.

Le second chapitre consiste en des estimations analytiques de sommes pondérées de valeurs de fonctions L de formes modulaires, dans l'esprit de techniques développées par Duke et Ellenberg. La motivation de départ d'un tel résultat est l'application de la méthode de Mazur dans le premier chapitre.

Le troisième chapitre est consacré à la recherche de généralisations de la méthode de Runge pour des variétés de dimension supérieure. Nous y redémontrons un résultat de Levin inspiré de cette méthode, avant d'en prouver une forme assouplie dite « de Runge tubulaire », plus largement applicable. Dans l'optique de recherche de points entiers de variétés modulaires, nous en donnons enfin un exemple d'utilisation à la réduction d'une surface abélienne en produit de courbes elliptiques.

Mots-clés Courbes elliptiques, courbes modulaires, représentations galoisiennes, problème d' uniformité de Serre, méthode de Mazur, théorèmes d'isogénie, formule des traces de Petersson, méthode de Runge, variétés modulaires de Siegel, fonctions thêta

Abstract

This thesis concerns the study of integral and rational points on some modular curves and varieties. After a brief introduction which describes the motivation and the setting of this topic as well as the main results of this thesis, the manuscript follows a threefold development.

The first chapter focuses on $\mathbb{Q}$-curves, and on the morphisms $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}_2(\mathbb{F}_p)$ that we can build with a $\mathbb{Q}$-curve for every prime p. We prove that, under good hypotheses, for p large enough with respect to the discriminant of the definition field of the $\mathbb{Q}$-curve, such a morphism is surjective, which solves a particular case of Serre's uniformity problem (still open in general). The main tools of the chapter are Mazur's method (based here on results of Ellenberg), Runge's method, and isogeny theorems, following the strategy of Bilu and Parent.

The second chapter covers analytic estimates of weighted sums of L-function values of modular forms, in the fashion of techniques designed by Duke and Ellenberg. The initial goal of such a result is the application of Mazur's method in the first chapter.

The third chapter is devoted to the search for generalisations of Runge's method for higher-dimensional varieties. Here we prove anew a result of Levin inspired by this method, before proving an enhanced version called « tubular Runge », more generally applicable. In the perspective of studying integral points of modular varieties, we finally give an example of application of this theorem to the reduction of an abelian surface in a product of elliptic curves.

Keywords Elliptic curves, modular curves, Galois representations, Serre's uniformity problem, Mazur's method, isogeny theorems, Petersson trace formula, Runge's method, Siegel modular varieties, theta functions

Remerciements

Tout d'abord, je tiens à remercier du fond du coeur mon directeur de thèse, Pierre Parent. Il m'a accordé une grande confiance très tôt, tout en sachant me remettre sur la bonne voie quand c'était nécessaire. J'ai beaucoup appris durant ces quatre ans et demi, et pour une grande part grâce au choix des sujets de la thèse. Il a également fourni dans les moments chauds une aide considérable, me faisant sans doute battre au passage le record de France de Skype dans les derniers mois.

Ensuite, je suis très honoré d'avoir eu comme rapporteurs Aaron Levin et René Schoof, deux personnes dont j'admire grandement le travail, et avec qui je serais enchanté de discuter plus longuement de mathématiques.

J'ai eu la chance de rencontrer avant ce jour tous les autres membres du jury, qui m'ont chacun apporté quelque chose au cours de la thèse sous forme de conseils avisés ou de discussions enrichissantes, et je les en remercie.

Ne faisant pas partie du jury, mais ayant également joué un rôle crucial pendant la thèse, je remercie très chaleureusement Gaël Rémond (entre autres, mais pas seulement, pour sa relecture minutieuse de mon premier article et son aide formidable sur les théorèmes d'isogénie), et Fabien Pazuki, pour sa bonne humeur permanente et les discussions enrichissantes sur les fonctions thêta.

Thésard expatrié à Lyon désormais, je n'ai pas oublié tout le personnel de l'IMB (administration comprise), grâce à qui l'atmosphère bordelaise fut si plaisante, et dont beaucoup m'ont donné un coup de main à un moment ou à un autre. Merci à vous.

Humainement, cette thèse restera aussi une aventure géniale grâce à l'ambiance entre doctorants de l'IMB. Quiconque nous aura vus faire les quatre cent coups aura noté mon plaisir d'avoir fréquenté Romain (grand amateur de houblon et de lecture devant l'éternel), Zoé (footix, membre fondatrice de la team, bretonne honoraire) et Bruno (érudit, fan de Neymar, Maître Pokémon et compagnon de thèse aujourd'hui). Merci également à Samuele, Pierre, Giovanni, Jocelyn, JB, Alan, Marie, Alice, Camille, Raphaël, Clément et aux deux Marc (qui ont échappé de justesse à un troll dans ce paragraphe).

Entre 2007 et 2011, c'est à l'ENS Lyon que j'eus la chance de découvrir les mathématiques, et plus précisément la théorie des nombres. En garder tant de bons souvenirs (notamment de l'année d'agrèg) et continuer à en construire avec eux est un gage de la qualité des amis que j'y ai rencontré : Julien (à jamais, je t'appellerai « dude »), Val (qui marchait dans la rue avec des chaussures de ski), Olivier « bras de singe » Glorieux, Sébastien (le Portugais), Régis Robert (« j'ai mal au dos mais c'est drôle »), Blanche (qu'il ne vaut mieux pas emmerder), Daniel (non, Guingamp n'est pas un grand d'Europe), ainsi que Mathieu, Richard, et Etienne.

Gaulois de nouveau depuis septembre 2014, j'ai rencontré de nouvelles personnes, à commencer par mon cobureau Loïc (Lorrain, supporter de l'OM et probabiliste, soyez gentils avec lui), dont le soutien moral et pratique pendant les derniers mois a été immense, j'espère pouvoir lui rendre la pareille. Notre ancien cobureau, Sylvain, est tout autant un modèle de gentillesse et de camaraderie, en plus d'être un sacré buteur. Pour compléter la bande, j'ajouterai Romain, Alexandre et Thomas, de solides compagnons de rigolade et de beauferie.

« As-tu envie de venir manger chez nous ce soir ? » Paul et Marguerite m'ont ainsi invité quasiment une fois par semaine depuis plus d'un an, si ma thèse était à la hauteur de leur convivialité elle ferait probablement trois volumes de plus. Pour les nombreuses soirées plus ou moins sobres (et le fait de supporter un matheux en leur sein), je remercie également mes amis lettreux, notamment Anaïs, Clémence et Txapu (ce dernier ayant d'ailleurs gentiment accepté de relire les pages

en anglais du manuscrit).

Mes deux petits frères, Nicolas et Victor, me connaissent sans doute mieux que moi-même (à quelle phrase de Kaamelott suis-je en train de penser en ce moment ?), donc merci pour s'être marrés avec moi toutes ces années. Quant à mes parents, j'ai pour eux toute la gratitude possible, pour avoir veillé sur moi depuis plus d'un quart de siècle, malgré mon intérêt plus que suspect pour tous ces trucs vachement abstraits (« c'est quoi ton titre de thèse déja ? »).

Enfin, pour ces six années passées ensemble, et ton support dans les bons comme les mauvais moments (dont le stress de la thèse), un grand merci à Céline, nous avons parcouru le chemin et j'espère bien qu'on va continuer, y compris pour ta propre soutenance l'année prochaine.

Table des matières

Introduction

Les équations diophantiennes sont un des sujets les plus riches et les plus anciens des mathématiques. Étudier une *équation diophantienne* (c'est-à-dire une équation polynomiale à coefficients entiers), c'est chercher à décrire ses solutions entières, rationnelles ou dans un corps de nombres. Par exemple et pour ne pas citer l'équation de Fermat, les équations

$$X^n + Y^n = Z^2, \quad \text{et} \quad X^n + Y^n = Z^3$$

étudiées dans [Poo98] n'ont, pour $n \geq 4$, aucune solution entière (x, y, z) sans facteur commun telle que $xyz \notin \{-1, 0, 1\}$. De même, on parle de *système d'équations diophantiennes* lorsqu'on a un système d'équations polynomiales dont on cherche de telles solutions. Un point de vue crucial pour ces systèmes d'équations diophantiennes est de les considérer comme constituant un objet géométrique (*schéma* pour le point de vue entier, ou *variété algébrique* pour le point de vue rationnel), de sorte que ses solutions entières (resp. rationnelles) sont les points à valeurs entières du schéma (resp. rationnelles de la variété algébrique). On espère alors pouvoir décrire les solutions entières ou rationnelles des équations en fonction de la structure géométrique, voire topologique, de l'objet associé.

Une illustration particulièrement frappante est donnée par le théorème de Mordell-Faltings, qui énonce que pour une bonne équation diophantienne (définissant une courbe algébrique projective lisse), on n'a qu'un nombre fini de solutions sur tout corps de nombres si la surface de Riemann associée est de genre au moins 2. Cette finitude n'épuise cependant pas la question car on s'intéresse plus précisément à l'existence de solutions.

Le thème de cette thèse sera le développement de diverses techniques de détermination (d'inexistence ou de finitude pour la plupart) des points entiers ou rationnels de certains schémas ou variétés algébriques.

Appliquées à certains espaces de modules, ces techniques permettent de répondre à des questions d'arithmétique. Une de ces questions est le problème d'uniformité de Serre, qui est la motivation de la première partie de cette thèse et que nous allons expliquer maintenant. Une *courbe elliptique E sur $\mathbb{Q}$* est une courbe algébrique qu'on peut définir par une certaine équation

$$E: \quad Y^2 = X^3 + aX + b \qquad a, b \in \mathbb{Q},$$

à laquelle on ajoute un point ∞, qui correspond à la solution $(0:1:0)$ de l'équation homogénéisée $Y^2Z = X^3 + aXZ^2 + bZ^3$ dans $\mathbb{P}^3$. On suppose de plus que $4a^3 + 27b^2 \neq 0$ pour assurer qu'elle est lisse.

Une telle courbe elliptique est canoniquement munie d'une loi de groupe abélien sur ses points à valeurs dans une extension fixée $K/\mathbb{Q}$, donnée par des fractions rationnelles à coefficients rationnels en fonction des coordonnées, et dont le point ∞ est l'élément neutre. On note $E(K)$ le groupe obtenu. Si K est un corps de nombres, le groupe abélien $E(K)$ est de type fini d'après le célèbre théorème de Mordell-Weil.

On s'intéresse plus particulièrement au groupe $E(K)_{\text{tors}}$ des points de torsion de $E(K)$. Celui-ci est fini, mais on souhaite comprendre quels peuvent être son exposant et sa structure. Étant donné la nature de la loi de groupe sur E, pour toute extension K de $\mathbb{Q}$, si $P = (x, y)$ est un point de $E(K)$ et σ un automorphisme de $K/\mathbb{Q}$, le point $\sigma(P) := (\sigma(x), \sigma(y))$ est encore un point de $E(K)$. De plus, l'application $\sigma : E(K) \to E(K)$ ainsi définie est un automorphisme de groupe. Pour

tout p premier, on note $E[p]$ l'ensemble des points de p-torsion de E à valeurs dans $\overline{\mathbb{Q}}$. L'action ci-dessus définit donc une *représentation galoisienne*

$$\rho_{E,p} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(E[p]).$$

De plus, $E[p]$ est isomorphe à $(\mathbb{Z}/p\mathbb{Z})^2$ et on en déduit donc une représentation de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ à valeurs dans $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$. Remarquons qu'on peut considérer $E[p]$ comme l'ensemble des solutions d'un système d'équations diophantiennes en (x, y) (auquel on ajoute ∞). Comprendre ce système d'équations diophantiennes, c'est en particulier comprendre quel est le plus petit corps de nombres $K_{E,p}$ pour lequel tous les points de $E[p]$ sont à valeurs dans $K_{E,p}$ (par exemple, est-il possible que $E(\mathbb{Q})[p]$ soit différent de $\{0\}$, voire $E[p]$ tout entier ?). La représentation $\rho_{E,p}$ permet d'attaquer ce problème, puisque son noyau définit exactement $K_{E,p}$ par correspondance de Galois. Ainsi, ce corps est grand si et seulement si l'image de $\rho_{E,p}$ est grande (c'est-à-dire que l'action de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ permute beaucoup les points de $E[p]$) : on parle alors de « grosse image galoisienne ».

On cherche donc à savoir à quel point l'image de $\rho_{E,p}$ est grosse. La réponse à cette question, apportée par Serre [Ser72], est que sauf dans un cas particulier (on dit que E est *à multiplication complexe*), dès que p est un nombre premier assez grand, la représentation $\rho_{E,p}$ permute autant les points de $E[p]$ que possible, c'est-à-dire qu'elle est surjective dans $\mathrm{GL}(E[p])$ (dans le cas de la multiplication complexe, l'image de $\rho_{E,p}$ est plus petite mais encore mieux comprise). Remarquons que, pour p assez grand, on sait donc que le système d'équations diophantiennes associé à $E[p]$ n'a aucune solution rationnelle mis à part le point ∞. En fait, nous montrons comme résultat intermédiaire (théorème I.7) dans la section I.8 une version explicite (et légèrement plus forte) de ce théorème de surjectivité relative, prouvant en particulier que $\rho_{E,p}$ est surjective dès que

$$p > 10^7 \left(\max(h_{\mathcal{F}}(E), 985)\right)^2, \tag{1}$$

où $h_{\mathcal{F}}(E)$ est la hauteur de Faltings stable de E (Formule (I.10) [1]). C'est à notre connaissance la première formulation totalement explicite en la hauteur de E. Rappelons que grossièrement, une fonction hauteur sur un ensemble d'objets de même nature est une mesure de la complexité de construction de ces objets, par exemple une hauteur naturelle sur $\mathbb{Z}$ est la valeur absolue.

Maintenant, on peut se demander si cette surjectivité pour p assez grand vient du choix des coefficients a et b définissant E, ou bien si le fait même d'être une courbe elliptique (sans multiplication complexe) sur $\mathbb{Q}$ entraîne automatiquement (et indépendamment de E) que $\rho_{E,p}$ est surjective pour p assez grand. Pour formuler rigoureusement cette question, existe-t-il une *borne uniforme* C telle que pour toute courbe elliptique E sur $\mathbb{Q}$ sans multiplication complexe, la représentation $\rho_{E,p}$ est surjective dès que $p > C$? C'est ce qu'on appelle le *problème d'uniformité de Serre*.

Pour y répondre, on découpe ce problème en quatre parties disjointes, correspondant à quatre types de sous-groupes stricts maximaux de $\mathrm{GL}_2(\mathbb{F}_p)$ (proposition I.1.5), dont le cas « exceptionnel » (section I.1.3) est déja traité dans [Ser72]. Par exemple, on dit que $\rho_{E,p}$ est dans le cas « Borel » si elle stabilise globalement un sous-groupe C_p cyclique d'ordre p de $E[p]$. Cette perspective permet de considérer les courbes elliptiques dans chacun des trois cas restants comme des points rationnels de certaines courbes algébriques, appelées les *courbes modulaires*. Ainsi, pour tout nombre premier p, la courbe modulaire $X_0(p)$ sur $\mathbb{Q}$ (paragraphe I.1.4) est telle qu'à part ses deux pointes, les points rationnels de $X_0(p)$ correspondent aux paires (E, C_p), où E est une courbe elliptique sur $\mathbb{Q}$ et C_p un sous-groupe cyclique d'ordre p de E comme ci-dessus.

On a donc réduit le problème d'uniformité de Serre sur $\mathbb{Q}$ à trois problèmes diophantiens distincts. Pour celui du cas « Borel », il s'agit de montrer que pour p assez grand, la courbe $X_0(p)$ n'a pas d'autres points rationnels que ses points *triviaux*, c'est-à-dire les deux pointes et les points rationnels éventuellement associés aux courbes elliptiques à multiplication complexe. C'est le cas pour $p > 37$ ([Maz78], Théorème 1). On peut en déduire que pour une courbe elliptique E sur $\mathbb{Q}$, le groupe $E(\mathbb{Q})_{\mathrm{tors}}$ est de cardinal au plus 12 ([Maz78], Théorème 2). Rappelons maintenant

1. Dans l'introduction, tout référence à une définition ou à un résultat de la thèse est destinée à éclaircir les notions de la phrase qui la précède, ou à en indiquer la preuve.

que le *j-invariant* d'une courbe elliptique E sur K est un invariant $j(E) \in K$ associé à E qui caractérise E à isomorphisme près. En tant que fonction sur une courbe modulaire, c'est une fonction rationnelle dont les pôles sont les pointes de cette courbe. Un point crucial de la preuve de ce premier théorème de [Maz78] est montrer que si (E, C_p) appartient à $X_0(p)(\mathbb{Q})$, alors il est en fait *entier pour le j-invariant*, c'est-à-dire que $j(E) \in \mathbb{Z}$. Ce point est obtenu par la *méthode de Mazur* (section I.2), conçue dans [Maz77]. Le célèbre théorème de Siegel sur les points entiers des courbes permet alors théoriquement de borner la hauteur des points de $X_0(p)(\mathbb{Q})$ différents des pointes, mais il ne suffit pas ici, et Mazur conclut la preuve avec un argument différent.

Dans le cas « normalisateur de Cartan déployé », le travail de [BP11a] permet d'obtenir que la courbe modulaire correspondante, notée $X_{\text{split}}(p)$ (paragraphe I.1.4), n'a pas de points rationnels non triviaux pour $p > 2 \cdot 10^{11}$. L'article complémentaire [BPR13] améliore cette borne en $p > 13$. Les outils employés sont la méthode de Mazur combinée à une technique de majoration de la hauteur des points entiers (la *méthode de Runge*, qui est le fil conducteur de cette thèse, et le thème central du chapitre III) et à une minoration de la hauteur de Faltings stable dans l'esprit de l'inégalité (1). Nous allons utiliser une structure de preuve similaire pour notre problème, énoncé ci-dessous.

Dans le dernier cas dit « normalisateur de Cartan non déployé », les courbes modulaires associées ne permettent pas, dans l'état actuel des connaissances, d'utiliser la méthode de Mazur (section I.6), et le problème d'uniformité de Serre pour les courbes elliptiques sur $\mathbb{Q}$ reste donc ouvert.

La première partie de cette thèse (chapitres I et II, qui reprennent les articles [LF] et [LF15]) est consacrée non pas aux courbes elliptiques sur $\mathbb{Q}$, mais à un autre type de courbes elliptiques : les $\mathbb{Q}$-*courbes* (définition I.1.1). On leur associe un *degré* $d(E) \geq 1$, qui mesure à quel point elles ne sont pas définies sur $\mathbb{Q}$, de sorte que $d(E) = 1$ si et seulement si E est définie sur $\mathbb{Q}$. Elles sont *strictes* si ce ne sont pas des quotients par un groupe fini de courbes elliptiques sur $\mathbb{Q}$. Pour tout p premier ne divisant pas $d(E)$, on a de manière analogue à $\rho_{E,p}$ une représentation $\mathbb{P}\overline{\rho}_{E,p}$ de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, qui est non plus à valeurs dans $\mathrm{GL}(E[p])$ mais dans $\mathrm{PGL}(E[p])$ (définition I.1.2). Pour toute $\mathbb{Q}$-courbe définie sur un corps quadratique K, on a encore une surjectivité relative explicite grâce au théorème I.7, et le problème d'uniformité de Serre se pose de la même manière, avec de nouveau un découpage en quatre cas pareillement nommés. L'objectif du chapitre I est de résoudre ce problème d'uniformité pour les $\mathbb{Q}$-courbes strictes définies sur un certain corps quadratique imaginaire K.

Pour les cas « Borel » et « normalisateur de Cartan déployé », la méthode de Mazur s'adapte aux $\mathbb{Q}$-courbes (sections I.2, I.4 et I.5), avec l'aide de quelques compléments techniques (section I.3).

Pour le cas « normalisateur de Cartan non déployé », une astuce d'Ellenberg (voir [Ell04] ou la section I.6) permet de voir que si K est quadratique imaginaire et E stricte, on peut sauver la méthode de Mazur, ce qui est surprenant au premier abord vu que cela est pour l'instant impossible sur $\mathbb{Q}$. Pour ce faire, on a besoin d'estimations analytiques de fonctions L de formes modulaires tordues par le caractère de Dirichlet χ_K associé à K, qui sont l'objet du chapitre II. Les calculs de la section II.3 prouvent alors que si p est assez grand en fonction de la valeur absolue D_K du discriminant de K, alors $j(E) \in \mathcal{O}_K$. Le théorème II.1 donne une version effective de ce résultat, raffinant les calculs originaux de [Ell04]. La section II.4 fournit l'énoncé analogue pour K quadratique réel (théorème II.2), qui est alors applicable à la $\mathbb{Q}$-courbe E seulement si $\chi_K(d(E)) = 1$ (remarque I.6.1).

Grâce à la méthode de Mazur ainsi appliquée, le problème de surjectivité uniforme pour de telles $\mathbb{Q}$-courbes de degré d devient un problème de points entiers quadratiques sur $X_0(d)$, dont on majore comme dans [BP11a] la hauteur par la méthode de Runge (section I.7), qui est notablement plus simple que son application à des familles générales de courbes modulaires (section III.3).

Enfin, on minore la hauteur de nos points entiers correspondant à des $\mathbb{Q}$-courbes telles que $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas surjective grâce à un avatar de la formule (1) plus haut (théorème I.7). Les bornes inférieure et supérieure sur la hauteur ainsi obtenues sont contradictoires pour p trop grand, d'où un résultat de surjectivité uniforme pour les $\mathbb{Q}$-courbes strictes sur un corps quadratique imaginaire.

Le résultat principal du chapitre I (rendu plus précis dans le théorème I.8) est alors le suivant, publié dans [LF]. Il s'agit d'un des premiers résultats connus de surjectivité uniforme explicite pour certaines familles infinies de courbes elliptiques sur des corps de nombres (une discussion sur la taille et la paramétrisation de ces familles se trouve dans l'exemple I.1.2).

Théorème (Surjectivité uniforme pour les $\mathbb{Q}$-courbes strictes)**.**
Soit K un corps quadratique imaginaire de discriminant $-D_K$ et E une $\mathbb{Q}$-courbe stricte définie sur K. Pour tout nombre premier p ne divisant pas $D_K d(E)$ et tel que

$$p > \max(2 \cdot 10^{13}, 50 D_K^{1/4} \log D_K),$$

la représentation $\mathbb{P}\overline{\rho}_{E,p}$ est surjective.

La stratégie de preuve esquissée ci-dessus illustre le rôle essentiel de la méthode de Runge pour majorer la hauteur des points entiers. C'est cette méthode que l'on se propose dans le chapitre III de cette thèse d'approfondir et de généraliser, pour l'appliquer à des variétés de dimension supérieure correspondant donc à des systèmes d'équations diophantiennes plus élaborés. Un des objectifs de ce chapitre est de pouvoir obtenir des résultats similaires à ceux de la méthode de Runge dans le chapitre I pour des variétés modulaires de Siegel, qui pourraient donc se reformuler en termes de propriétés de variétés abéliennes. La section III.1 fournit les notions et le vocabulaire nécessaires pour l'exposition des énoncés du chapitre. La section III.2 explique les résultats théoriques derrière la méthode de Runge pour toute courbe algébrique, puis comment exécuter la méthode à partir de ceux-ci. Pour en donner la saveur, soit C une courbe algébrique projective lisse définie sur un corps de nombres K et ϕ une fonction K-rationnelle non constante sur C. On note, pour toute extension L de K, r_L le nombre d'orbites des pointes de ϕ par $\mathrm{Gal}(\overline{K}/L)$. Alors, l'ensemble des points $P \in C(L)$ tels que $\phi(P) \in \mathcal{O}_{L,S_L}$ pour une certaine paire (L, S_L) avec S_L un ensemble de places de L contenant les places archimédiennes et tel que

$$|S_L| < r_L \tag{2}$$

est fini (théorème III.2). Cette condition sur (L, S_L) est la *condition de Runge*.

L'idée derrière cette condition est la suivante. On prend $P \in C(L)$ tel que $\phi(P) \in \mathcal{O}_{L,S_L}$, et on cherche à borner la taille d'une fonction auxiliaire $|\psi(P)|_v$ en chaque place v de L. On peut en fait construire des fonctions auxiliaires $\psi_1, \cdots, \psi_{r_L}$ de $L(C)$ indexées par les orbites de pôles de ϕ par $\mathrm{Gal}(\overline{K}/L)$ de telle sorte que chacune a pour ensemble de pôles une telle orbite. Il suffit de borner la taille d'un $|\psi_i(P)|_v$ pour chaque place v de S_L, car pour les places $v \notin S_L$, $|\phi(P)|_v \leq 1$ donc les $|\psi_i(P)|_v$ le sont. Alors, $|\psi_i(P)|_v$ est grand si et seulement si P est v-adiquement près d'un des pôles de ψ_i, et comme les ensembles des pôles des ψ_i sont distincts entre eux, on ne peut pas être trop près de deux d'entre eux en même temps. Cela signifie que pour chaque place v de S_L, il existe au plus un indice i tel que $|\psi_i(P)|_v$ est grand. On élimine si nécessaire un indice par place de S_L, et par la condition de Runge, il reste un indice i tel que $|\psi_i(P)|_v$ est petit pour tout $v \in S_L$. En conséquence, la hauteur de $\psi_i(P)$ est petite et P appartient à un ensemble fini (proposition III.1.1). En détaillant cette preuve, on peut même s'assurer que cette construction est uniforme en (L, S_L), et obtenir une estimation sur la hauteur associée à ϕ et pas seulement en associée à ψ_i (théorème III.3).

La section III.3 reprend et généralise l'application de la méthode de Runge aux courbes modulaires, d'après [BP11b]. La section III.4 est consacrée à l'obtention de généralisations en dimension supérieure. De substantielles difficultés apparaissent par rapport au cas des courbes. La première est que pour parler de points entiers, il faut avoir défini un bord composé d'un certain nombre de diviseurs effectifs réduits, dans nous verrons qu'ils doivent avoir certaines propriétés géométriques, par exemple être *amples* (ce qui était automatique dans le cas des courbes mais peut être une hypothèse très forte dans certaines variétés). Énonçons maintenant notre résultat « à la Runge » en toute dimension. On désigne par (L, S_L) une paire avec L un corps de nombres et S_L un ensemble de places de L contenant les places archimédiennes. Soit X une variété normale projective sur un corps de nombres K et $D_1, \cdots, D_r$ des diviseurs amples effectifs réduits de X. On choisit $\mathcal{X}$ un

modèle normal propre de X sur $\mathcal{O}_K$ et $\mathcal{D}_1, \cdots, \mathcal{D}_r$ leur adhérence de Zariski dans $\mathcal{X}$, d'union $\mathcal{D}$. Contrairement au cas des courbes, les diviseurs $D_1, \cdots, D_r$ peuvent s'intersecter mutuellement, et donc un point peut être v-adiquement proche de plusieurs d'entre eux à la fois. Le raisonnement derrière la méthode de Runge (un principe des tiroirs éliminant place par place les diviseurs trop près de notre point, dans le but qu'il en reste un) nous oblige donc cette fois-ci à éliminer des diviseurs paquets par paquets quand ils s'intersectent, et la condition de Runge (2) se détériore alors en

$$m|S_L| < r, \tag{3}$$

avec m l'entier tel que l'intersection de $(m+1)$ diviseurs distincts parmi les $D_1, \cdots, D_r$ doit être vide (*condition de Runge générale*). Ceci est énoncé précisément et en plus grande généralité dans le théorème III.7, qui est une reformulation du théorème 4 de [Lev08] dans les cas amples et gros.

Notons qu'un avantage majeur de la méthode de Runge est que lorsque elle s'applique (et donne donc un résultat de finitude de points (L, S_L)-entiers pour certaines paires (L, S_L) qui vérifient la condition), elle s'applique également à la réunion de tous les points (L, S_L)-entiers considérés, qui doit également être finie (afin de mettre ceci en valeur, nous avons refait la preuve du théorème III.7 dans la section III.4). En un sens, elle ne différencie pas les places de non-intégralité, et est sensible seulement à leur nombre. Ceci n'était pas visible pour $X_0(p)$ dans la section I.7, car elle ne pouvait s'appliquer qu'à la place archimédienne infinie du corps quadratique imaginaire K, mais a été remarqué initialement par Bombieri [Bom83] pour les courbes. Cette dépendance seulement en le nombre de mauvaises places (et la finitude en groupant les points vérifiant la condition) constitue un des avantages majeurs de la méthode de Runge par rapport à des méthodes plus connues de majoration de hauteur de points entiers, comme la méthode de Baker par exemple.

Cependant, si le nombre d'intersections maximal m de l'équation (3) se trouve être trop grand, la condition de Runge générale n'est jamais satisfaite. Il apparaît alors une deuxième difficulté consubstantielle à la dimension supérieure : des diviseurs amples s'intersectent nécessairement, contrairement au cas des courbes puisque pour des points, être distincts et être disjoints sont synonymes... Plus précisément, si on a r diviseurs amples effectifs sur une variété algébrique lisse projective de dimension d, par des propriétés de l'intersection et l'hypothèse d'amplitude, si ces diviseurs sont en position générale, leur intersection totale est automatiquement non vide dès que $r \leq d$. Cette obstruction de l'intersection multiple (ainsi que le fait que les diviseurs impliqués peuvent très bien être fortement concourants, indépendamment de leur nombre et de leur géométrie) risque d'empêcher l'utilisation du théorème de Runge en dimension supérieure dans un certain nombre de cas. Nous avons donc développé une nouvelle notion de « Runge tubulaire », plus souple que l'énoncé classique du théorème 4 de [Lev08]. Elle se formule en fait comme un théorème de *concentration des points entiers*, que nous allons énoncer de manière simplifiée ci-dessous.

Avec les notations ci-dessus, soit Y un fermé de X et $\mathcal{Y}$ son adhérence de Zariski dans $\mathcal{X}$, qu'on va considérer comme un domaine supplémentaire à exclure. On note $M_{\overline{K}}$ les places de $\overline{K}$ et $\mathcal{U} = (U_v)_{v \in M_{\overline{K}}}$ une famille de voisinages respectifs de $Y(\overline{K}_v)$ dans la topologie v-adique de $X(\overline{K}_v)$ (définition III.1.4) qui doivent être en un sens « uniformément assez larges autour de Y » (pour la signification précise, cf. définition III.4.5). On peut penser, si Y est une courbe, à chaque voisinage U_v comme à un tube autour de cette courbe dans X, d'où le nom de *voisinage tubulaire de Y* qu'on donne à une telle famille $\mathcal{U}$ (remarque III.4.5). On dit qu'un point $P \in X(\overline{K})$ appartient à ce voisinage tubulaire si et seulement il appartient à un U_v pour un certain $v \in M_{\overline{K}}$. Notre théorème se formule alors de la manière suivante.

Théorème (Runge tubulaire)**.** *Avec les notations ci-dessus, soit $\mathcal{E}(L, S_L)(\mathcal{U})$ l'ensemble des points de $(\mathcal{X} \backslash \mathcal{D})(\mathcal{O}_{L,S_L})$ qui vérifient en plus la propriété d'être v-adiquement loin de Y en toute place de $M_{\overline{K}}$, c'est-à-dire n'appartenant pas à $\mathcal{U}$. Alors, pour tout voisinage tubulaire $\mathcal{U}$, la réunion de tous les $\mathcal{E}(L, S_L)(\mathcal{U})$ est finie, lorsque (L, S_L) parcourt les paires vérifiant la condition de Runge tubulaire*

$$m_Y|S_L| < r, \tag{4}$$

où m_Y est le plus petit entier tels que l'intersection de n'importe quels (m_Y+1) diviseurs distincts $D_1, \cdots, D_r$ est incluse dans Y (et non plus vide comme pour la condition de Runge générale).

Il est à noter que si $\mathcal{Y}$ est inclus dans $\mathcal{D}$ (ce qui sera le cadre que nous avons en tête), on sait que les $P \in (\mathcal{X}\backslash\mathcal{D})(\mathcal{O}_{L,S_L})$ ne sont pas v-adiquement trop près de Y pour $v \in M_L\backslash S_L$. La condition ainsi exposée complète donc l'hypothèse d'intégralité en les places de S_L, en particulier en toutes les places archimédiennes.

Ce résultat est bien un énoncé de concentration, car il signifie que pour tout voisinage tubulaire $\mathcal{U} = (U_v)_{v\in M_{\overline{K}}}$, à un nombre fini de points près, les $P \in (\mathcal{X}\backslash\mathcal{D})(\mathcal{O}_{L,S_L})$ où (L, S_L) vérifie la condition de Runge se *concentrent* près de Y, c'est-à-dire appartiennent au voisinage tubulaire $\mathcal{U}$. L'énoncé complet (plus général et détaillé) du théorème de Runge tubulaire constitue le théorème III.8, et la preuve de celui-ci est le but de la section III.4.

Remarquons que ce théorème est automatique lorsque Y est ample, au même titre que la propriété de Northcott pour la hauteur associée à un diviseur ample (proposition III.4.6). À l'opposé, lorsque Y est vide, c'est exactement le théorème III.7. Pour appliquer ce théorème hybride, il faut donc faire un compromis pour parvenir à avoir à la fois Y et m_Y les plus petits possibles.

Ce théorème peut également être comparé à ceux de [CLZ09], notamment au théorème d'Autissier (pour les surfaces) et au théorème CLZ (Corvaja, Levin et Zannier) pour les variétés de dimension au moins 3. Ces deux derniers sont de saveur similaire à notre théorème de concentration, mais voici quelques différences notables. Tout d'abord, ils ne demandent pas de condition sur la taille de l'ensemble de places S_L, pas de condition aussi stricte d'intersection que dans notre théorème, et pas non plus d'hypothèse supplémentaire d'exclusion d'un voisinage tubulaire (qui, comme on l'a vu, est surtout une vraie condition supplémentaire pour les places de non-intégralité et les places archimédiennes). En revanche, la finitude énoncée (bien que le domaine d'exclusion Y de ces théorèmes soit absolu) dépend de l'ensemble de places S_L (en particulier, il ne dit pas si la réunion de tous ces ensembles finis est finie ou non), et on ne sait rien dans le cas des diviseurs gros. Enfin, l'outil de base de ces résultats étant le théorème du sous-espace de Schmidt, il paraît difficile de les rendre effectifs, contrairement à notre théorème.

L'objectif de la section III.5 est une application de notre théorème de Runge tubulaire aux variétés modulaires de Siegel (théorème III.10) dont nous allons ici énoncer le cas particulier en niveau $n = 2$ (théorème III.9), qui est plus parlant. La variété $A_2(2)_{\mathbb{C}}$ est la variété modulaire de Siegel associée aux surfaces abéliennes principalement polarisées munies d'une structure symplectique de niveau 2 (définition-proposition III.5.3), et $A_2(2)^S_{\mathbb{C}}$ est sa compactification de Satake (définition III.5.6). Ces deux variétés admettent des modèles normaux propres naturels sur $\mathbb{Z}[1/2]$ (définition-proposition III.5.14).

Dans ce contexte, le « bord » (lieu par rapport auquel on définit l'intégralité) a une bonne interprétation modulaire comme lieu de dégénérescence de surfaces abéliennes. En effet, il se compose du « bord fort », lieu des surfaces semi-abéliennes avec partie torique non nulle, et du « bord faible », lieu des surfaces abéliennes principalement polarisées non simples, dont les composantes constituent les diviseurs effectifs définissant l'intégralité usuelle pour le théorème.

Cette situation est analogue au cas des courbes modulaires, où le bord correspond au lieu des courbes elliptiques dégénérées, c'est-à-dire aux pointes, et la distinction entre bord fort et bord faible n'y a pas de sens.

Théorème (Runge pour les produits de courbes elliptiques sur $\mathcal{A}_2(2)^S$)**.**

Pour un voisinage ouvert U du bord fort $\partial A_2(2)_{\mathbb{C}} := A_2(2)^S_{\mathbb{C}}\backslash A_2(2)_{\mathbb{C}}$ pour la topologie complexe usuelle, soit $\mathcal{E}(U)$ l'ensemble des points $P \in A_2(2)(\overline{\mathbb{Q}})$ représentant le triplet (A, λ, α_2) (à isomorphisme à extension des scalaires près), avec A une surface abélienne, λ une polarisation principale et α_2 une structure symplectique de niveau 2 sur A tel que :

- *Le point P est à valeurs dans une certaine extension finie L de $\mathbb{Q}$.*
- *La variété abélienne A a potentiellement bonne réduction en toute place finie, et quelle que soit l'extension des scalaires de $A_2(2)(L)$ à $A_2(2)_{\mathbb{C}}$, l'image du point P n'appartient pas à l'ouvert U.*
- *Le nombre s_L de places v de L telles que*
  - *v est archimédienne*
  - *v divise 2*

– *La réduction modulo v de (A, λ) dégénère à extension des scalaires près en un produit de courbes elliptiques (c'est-à-dire appartient au bord faible)*

vérifie la condition

$$s_L < 10.$$

Alors, l'ensemble $\mathcal{E}(U)$ est fini pour tout ouvert U contenant $\partial A_2(2)_{\mathbb{C}}$.

Dans cette application du théorème, on a utilisé un voisinage tubulaire naturel de $\partial \mathcal{A}_2(2)_{\mathbb{Q}}$. En effet, pour tout place finie $v \in M_{\overline{\mathbb{Q}}}$ associée à un certain idéal premier $\mathfrak{P}$, on peut considérer comme bord U_v l'ensemble des points $P_v \in A_2(2)^S(\overline{\mathbb{Q}}_v)$ dont la réduction modulo $\mathfrak{P}$ appartient à $\partial A_2(2)_{\mathfrak{P}}$. Si P_v représente une surface abélienne A_v, cela signifie que celle-ci n'a pas potentiellement bonne réduction modulo $\mathfrak{P}$, par construction des schémas $\mathcal{A}_2(2)$ et $\mathcal{A}_2(2)^S$ (et on peut sauver cette vision pour v au-dessus de 2 malgré que $\mathcal{A}_2(2)$ ne soit pas construit en caractéristique 2). Il reste donc simplement à compléter cette famille d'ouverts avec des ouverts de $v \in M_{\overline{\mathbb{Q}}}$ où v est archimédienne, et pour cela un moyen simple est de prendre un unique ouvert U de $A_2(2)^S_{\mathbb{C}}$ contenant $\partial A_2(2)_{\mathbb{C}}$, d'où la formulation ci-dessus. La condition de Runge $s_L < 10$ provient du fait que les diviseurs en jeu sont au nombre de dix, amples et effectifs sur $A_2(2)^S_{\mathbb{Q}}$ et d'intersection vide hors de $\partial A_2(2)_{\mathbb{Q}}$ (définition-proposition III.5.29 (b)), donc $m_{\partial A_2(2)} = 1$.

Notons enfin que nous donnons également dans les paragraphes III.5.1 et III.5.2 des descriptions de bonnes compactifications toroïdales de certains espaces de modules, ainsi que de la géométrie de leurs diviseurs, comme travail préparatoire à des applications futures de notre théorème de Runge tubulaire à d'autres variétés modulaires de Siegel, et d'autres diviseurs.

Introduction

Diophantine equations are among the most ancient and noble fields of mathematics. To study *diophantine equations* (that is, polynomial equations with integer or rational coefficients), is to study their integral or rational solutions (or more generally, their solutions in a number field). As an example, and to avoid quoting the famous Fermat equation, the equations

$$X^n + Y^n = Z^2, \quad \text{and} \quad X^n + Y^n = Z^3$$

studied in [Poo98] have, for $n \geq 4$ no integral solution (x, y, z) without commmon factor such that $xyz \notin \{-1, 0, 1\}$. Likewise, we call *system of diophantine equations* a system of polynomial equations for which we search for such solutions. A crucial viewpoint for these diophantine equations (or systems of) is to consider them as making up a geometric object (called *scheme* for the integral viewpoint, or *algebraic variety* for the rational one), such that its integral (resp. rational) solutions are the integer-valued (resp. rational-valued) points of the scheme (resp. algebraic variety). Our hope is then to be able to describe the integral or rational solutions of the equations with help of the geometrical (or even topological) structure of the corresponding object.

A particularly striking example of this is Mordell-Faltings' theorem, which states that for every " good " diophantine equation (defining a smooth algebraic projective curve), there is only a finite number of solutions on any number field if the corresponding Riemann surface has genus at least 2. This finiteness result does not solve everything though, because we are often interested in the problem of existence of solutions.

The theme of this thesis will be the construction of diverse tools of determination (of inexistence or finiteness, for most of them) of integral or rational points of some schemes or algebraic varieties.

Applied to moduli spaces, these techniques enable us to answer arithmetical questions. One of these is Serre's uniformity problem, which motivates the first part of this thesis and will be explained now. An *elliptic curve E over $\mathbb{Q}$* is an algebraic curve defined by some equation

$$E: \quad Y^2 = X^3 + aX + b \qquad a, b \in \mathbb{Q},$$

to which we add a point ∞, coming from the solution $(0 : 1 : 0)$ of the homogeneized equation $Y^2Z = X^3 + aXZ^2 + bZ^3$ in $\mathbb{P}^3$. We furthermore suppose that $4a^3 + 27b^2 \neq 0$ to ensure E is smooth.

Such an elliptic curve is canonically endowed with an abelian group law on its points in any extension K of $\mathbb{Q}$, given by rational fractions with rational coefficients in the coordinates, and whose point ∞ is the neutral element. We denote by $E(K)$ this abelian group, and if K is a number field, it is of finite type according to the famous Mordell-Weil theorem.

We focus here on the group $E(K)_{\text{tors}}$ of torsion points of $E(K)$. It is finite, but we wish to understand what can be its exponents and structure. Given the nature of the group law on $E(K)$ for any extension K, if $P = (x, y)$ belongs to $E(K)$ and σ is an automorphism of $K/\mathbb{Q}$, the point $\sigma(P) := (\sigma(x), \sigma(y))$ also belongs to $E(K)$, and the map $\sigma : E(K) \to E(K)$ thus defined is a group automorphism. For any prime p, we denote by $E[p]$ the set of p-torsion points of E in $\overline{\mathbb{Q}}$. The action above hence defines a *Galois representation*

$$\rho_{E,p} : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}(E[p]).$$

As $E[p]$ is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^2$, we obtain a representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$. Notice we can consider $E[p]$ as the set of solutions of a system of diophantine equations in (x, y) (to which we add ∞). To understand this system, we need to understand in particular what is the smallest number field $K_{E,p}$ for which all points of $E[p]$ have values in $K_{E,p}$ (e.g. can $E(\mathbb{Q})[p]$ be different from $\{0\}$, or even equal to $E[p]$?). The representation $\rho_{E,p}$ allows to tackle this question, since its kernel is exactly $K_{E,p}$ by Galois correspondence. Thus, this field is big if and only if the image of $\rho_{E,p}$ is big (that is, $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ permutes a lot the points of $E[p]$) : we call this phenomenon " big Galois image ".

Hence, we want to know how big is the image of $\rho_{E,p}$. The answer to this question, brought by Serre [Ser72], is that unless E has a special structure (that is, E has *complex multiplication*), for large enough p, the representation $\rho_{E,p}$ permutes as much as possible the points of $E[p]$, that is $\rho_{E,p}$ is surjective in $\mathrm{GL}(E[p])$ (for the complex multiplication case, the image of $\rho_{E,p}$ is smaller but better understood). Let us notice that for large enough p, we thus know that the system of diophantine equations associated to $E[p]$ has no rational solution except ∞. Actually, we obtain in the course of our proof (Theorem I.7) in section I.8 an explicit (and slighly stronger) version of Serre's surjectivity theorem, which implies in particular that $\rho_{E,p}$ is surjective when

$$p > 10^7 \left(\max(h_{\mathcal{F}}(E), 985)\right)^2, \tag{5}$$

with $h_{\mathcal{F}}(E)$ the stable Faltings height of E (Formula (I.10)) [2]. To our knowledge, this is the first surjectivity result which is completely explicit in the Faltings height of E. Let us recall that a height function is roughly a function on a set of objects of same nature measuring the complexity of construction of these objects (for example, a natural height on $\mathbb{Z}$ is the absolute value).

We can now ask ourselves if this surjectivity for large enough p comes from the choice of coefficients a and b defining E, or if the mere fact of being an elliptic curve (without complex multiplication) on $\mathbb{Q}$ automatically implies (independently of E) that $\rho_{E,p}$ is surjective for large enough p. To state rigorously this question, is there a *uniform bound* C such that for every elliptic curve E on $\mathbb{Q}$ without complex multiplication, the representation $\rho_{E,p}$ is surjective when $p > C$? This question is called *Serre's uniformity problem*.

To answer this, we split this problem into four disjoint parts, corresponding to four types of strict maximal subgroups of $\mathrm{GL}_2(\mathbb{F}_p)$ (proposition I.1.5), whose the " exceptional case " (section I.1.3) is already done in [Ser72]. We say for example that $\rho_{E,p}$ is in the " Borel " case if it globally stabilises a cyclic subgroup C_p of order p of $E[p]$. This perspective allows us to consider the elliptic curves in each remaining case as rational points on some algebraic curves, called *modular curves*. For every prime number p, the modular curve $X_0(p)$ on $\mathbb{Q}$ (paragraph I.1.4) is for example such that apart its two cusps, the rational points of $X_0(p)$ correspond to pairs (E, C_p), with E an elliptic curve on $\mathbb{Q}$ and C_p a cyclic subgroup of order p of E stable by $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

We have therefore reduced Serre's uniformity problem on $\mathbb{Q}$ to three distinct diophantine problems. For the " Borel " case, we have to prove that for large enough p, the curve $X_0(p)$ has only *trivial* rational points, that is its two cusps and the possible rational points coming from elliptic curves with complex multiplication. That is true for $p > 37$ ([Maz78], Theorem 1). We can infer from this that for any elliptic curve E over $\mathbb{Q}$, the group $E(\mathbb{Q})_{\mathrm{tors}}$ has order at most 12 ([Maz78], Theorem 2). Recall now that the *j-invariant* of an elliptic curve E over K is an invariant $j(E) \in K$ associated to E, characterising E up to isomorphism. As a function on a modular curve, it is a rational function whose poles are the modular curve's cusps. A crucial ingredient of the proof of the first theorem of [Maz78] is that if (E, C_p) belongs to $X_0(p)(\mathbb{Q})$, then it actually is *integral for the j-invariant*, that is $j(E) \in \mathbb{Z}$. This point is obtained by *Mazur's method* (section I.2), designed in [Maz77]. Consequently, the famous Siegel's theorem on integral points of curves theoretically allows us to bound the height of noncuspidal points of $X_0(p)(\mathbb{Q})$, but it is not sufficient here and Mazur ends the proof with a different argument.

In the " normalizer of split Cartan " case, [BP11a] proves that the associated modular curve, denoted by $X_{\mathrm{split}}(p)$ (paragraph I.1.4), has no nontrivial rational points for $p > 2 \cdot 10^{11}$. The

2. In this Introduction, every reference to a definition or a result of this thesis is used to enlighten the notions of the preceding sentence, or point towards its proof.

following article [BPR13] improves this bound to $p > 13$. The tools of these proofs are Mazur's method, a technique to bound above heights of integral points called *Runge's method* (which is the unifying thread of this thesis, and the central theme of Chapter III) and isogeny theorems to bound below the stable Faltings height in the fashion of inequality (5). We will use a similar proof structure for our problem, which is stated below.

In the last case, called « normalizer of nonsplit Cartan », the associated modular curves do not allow us, to our knowledge, to use Mazur's method (section I.6), thus Serre's uniformity problem remains open for elliptic curves over $\mathbb{Q}$.

The first part of this thesis (Chapters I and II, reproducing articles [LF] and [LF15]) is not focused on elliptic curves over $\mathbb{Q}$, but on another kind of elliptic curves called $\mathbb{Q}$-*curves* (Definition I.1.1). To each $\mathbb{Q}$-curve E we can associate a *degree* $d(E) \geq 1$, measure how far E is from being defined over $\mathbb{Q}$, so that $d(E) = 1$ if and only if E is defined over $\mathbb{Q}$. E is called *strict* if it is not isogenous (over $\overline{\mathbb{Q}}$) to an elliptic curve defined over $\mathbb{Q}$. For any prime p not dividing $d(E)$, we have, analogously to $\rho_{E,p}$, a representation $\mathbb{P}\overline{\rho}_{E,p}$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, not anymore in $\mathrm{GL}(E[p])$ but in $\mathrm{PGL}(E[p])$ (Definition I.1.2). For every $\mathbb{Q}$-curve over a quadratic field K, there is again a relative explicit surjectivity thanks to Theorem I.7, and Serre's uniformity problem is raised in the same fashion, and separated in four cases named in the same way. The goal of Chapter I is to solve this uniformity problem for strict $\mathbb{Q}$-curves over an imaginary quadratic field K.

For "Borel" and "normalizer of split Cartan" cases, Mazur's method can be adapted to $\mathbb{Q}$-curves (sections I.2, I.4 and I.5), with help of some technical complements (section I.3).

For the "normalizer of nonsplit Cartan" case, a trick of Ellenberg (see [Ell04] or section I.6) allows us to see that if K is imaginary quadratic and E strict, Mazur's method can be salvaged, which is surprising at first given it is impossible on $\mathbb{Q}$ for now. To do this, we need analytic estimates of L-functions of modular forms twisted by the Dirichlet character χ_K associated to K, which are the topic of Chapter II. The computations of section II.3 prove that if p is large enough compared to the discriminant of K, then $j(E) \in \mathcal{O}_K$. The theorem II.1 gives an effective version of this result, refining previous computations of [Ell04]. The section II.4 gives the analogous statement for K real quadratic (Theorem II.2), which is then applicable to the $\mathbb{Q}$-curve E only if $\chi_K(d(E)) = 1$ (Remark I.6.1).

Thanks to Mazur's method applied in this fashion, Serre's uniformity problem for such $\mathbb{Q}$-curves of degree d becomes a problem on integral quadratic points on $X_0(d)$, and we bound above their height, as in [BP11a], by Runge's method (section I.7), which is significantly simpler than its application to general families of modular curves (section III.3).

At last, we bound below our integral points' height using an avatar of formula (1) above (Theorem I.7). The lower and upper bounds on the height we obtain contradict each other for large enough p, thus giving a uniform surjectivity result on strict $\mathbb{Q}$-curves over an imaginary quadratic field. The main result of Chapter I (made more precise in Theorem I.8) is the following, published in [LF]. It is one of the first known results of explicit uniform surjectivity for some infinite families of elliptic curves on number fields (for a discussion about the size and parametrization of these families, see Example I.1.2).

Theorem (Uniform surjectivity for strict $\mathbb{Q}$-curves)**.**

Let K be an imaginary quadratic field of discriminant $-D_K$ and E a strict $\mathbb{Q}$-curve over K. For any prime p not dividing $D_K d(E)$ and such that

$$p > \max(2 \cdot 10^{13}, 50 D_K^{1/4} \log D_K),$$

the representation $\mathbb{P}\overline{\rho}_{E,p}$ is surjective.

The strategy of proof sketched above emphasizes the essential role of Runge's method to give an upper bound on the height of integral points. This is the method we want to develop and generalise in Chapter II, to apply it to higher-dimensional varieties therefore corresponding to more elaborate systems of diophantine equations. One of the goals of this chapter is to obtain results similar to Runge's method in Chapter I, but for Siegel modular varieties, which could be formulated in terms of properties of abelian varieties. The section III.1 brings up the necessary

notions and vocabulary for the results of the chapter. The section III.2 explains the theoretical results behind Runge's method for algebraic curves, and how to enforce the method behind the theory. To give a taste of these, let C be a smooth projective algebraic curve over a number field K and ϕ be a nonconstant K-rational function on C. For any extension L of K, we denote by r_L the number of orbits of poles of ϕ by $\mathrm{Gal}(\overline{K}/L)$. The set of points $P \in C(L)$ such that $\phi(P) \in \mathcal{O}_{L,S_L}$ for some pair (L, S_L) with S_L a set of places of L (including all the archimedean ones) and such that

$$|S_L| < r_L \tag{6}$$

is then finite (Theorem III.2). This condition on (L, S_L) is the *Runge condition*.

The idea behind this result is the following. Take $P \in C(L)$ such that $\phi(P) \in \mathcal{O}_{L,S_L}$, we want to bound above the absolue value of an auxiliary function $|\psi(P)|_v$ at every place v of L. Actually, we can build auxiliary functions $\psi_1, \cdots, \psi_{r_L}$ of $L(C)$ indexed by the orbits of poles of ϕ by $\mathrm{Gal}(\overline{K}/L)$, such that each one of this function has its poles constituting such an orbit. We only have to bound above $|\psi_i(P)|_v$ for some i and every place v of S_L, because for $v \notin S_L$, each $|\psi_i(P)|_v$ is small because $|\phi(P)|_v$ is. As $|\psi_i(P)|_v$ is large if and only if P is v-adically close to one of the poles of ψ_i, and the sets of poles of ψ_i are disjoint to each other, one cannot be too close to two of them at the same time. This means that for each place v of S_L, there is at most one index i such that $|\psi_i(P)|_v$ is large. If necessary, we put aside one index by place of S_L, and by Runge's condition, there remains one index i such that $|\psi_i(P)|_v$ is small for each $v \in S_L$. Consequently, the height of $\psi_i(P)$ is small and P belongs to a finite set (proposition III.1.1). When detailing this proof, we can even ensure that this process can be made uniform in (L, S_L), and obtain an estimate on the height associated to ϕ and not only to some auxiliary function ψ_i (Theorem III.3).

The section III.3 mentions once more and generalises the application of Runge's method to modular curves, after [BP11b]. The section III.4 is devoted to higher-dimensional generalisations of Runge's method. This brings about substantial difficulties compared to the case of curves. The first one is that to talk about integral points, we need to define a boundary composed of reduced effective divisors, whose as we will see must have some geometric properties, for example be *ample* (which was automatic in the case of curves but can be a very strong hypothesis on some varieties). Let us now state our result "à la Runge" in every dimension. We denote by (L, S_L) a pair with L a number field and S_L a set of places of L containing the archimedean ones. Let X be a normal projective variety over a number field K and $D_1, \cdots, D_r$ some reduced ample effective divisors of X. We choose $\mathcal{X}$ a proper normal model of X on $\mathcal{O}_K$ and $\mathcal{D}_1, \cdots, \mathcal{D}_r$ their Zariski closure in $\mathcal{X}$, whose union we denote by $\mathcal{D}$. It is important to note that as opposed to the case of curves, the divisors $D_1, \cdots, D_r$ can have nontrivial mutual intersection, therefore a point can be v-adically close of many of them simultaneously. The argument behind Runge's method (elimination place by place of the divisors too close of our point, made such that there remains one at the end) then compels us to eliminate divisors packs by packs when they intersect, therefore Runge condition (6) is altered to

$$m|S_L| < r, \tag{7}$$

with m the integer such that the intersection of $(m+1)$ distinct divisors amongst $D_1, \cdots, D_r$ must be empty (*general Runge condition*). This is stated precisely and with greater generality in Theorem III.7, which reformulates Theorem 4 of [Lev08] in the "ample" and "big" cases.

Notice that when Runge's method applies (and gives a finiteness result of (L, S_L)-integral points for some pairs (L, S_L) satisyfing Runge condition), it also applies and gives finiteness of the union of all (L, S_L)-integral points considered (to emphasize this, we have rewritten the proof of Theorem III.7 in section III.4). We can somewhat say that Runge's method does not distinguish the places of non-integrality and is only sensitive to their number. This was not visible for $X_0(p)$ in section I.7 as it was only appliable then to the unique archimedean place of the imaginary quadratic field K (although we could already see the bound was uniform in the choice of quadratic field), but has been first pointed out by Bombier [Bom83] for curves. This dependence only in the number of bad places (and the finiteness, even after putting together the points satisfying the condition) is one of the major assets of Runge's method, compared to more famous methods to bound heights of integral points, such as Baker's method for example.

On the other hand, if the maximal intersections number m of equation (7) happens to be too large, the general Runge condition is never satisfied. This is where a second difficulty intrinsic to higher dimension appears : ample divisors necessarily intersect, as opposed to the case of curves because distinct points are of course disjoint. More precisely, if we have r ample effective divisors on a smooth projective algebraic variety of dimension d, by ampleness and properties of intersection, if these divisors are in general position, their total intersection (as a set) is nonempty as soon as $r \leq d$. This obstruction of multiple intersection (as well as the fact that the considered divisors might very well intersect a lot, regardless of their number and geometry) threatens to prevent the application of Runge theorem in higher dimension in a number of cases. We have therefore developed a new notion of "tubular Runge", more flexible than Theorem 4 of [Lev08]. This result is actually a statement of *concentration of integral points*, that we will give in a simplified version below.

With the previous notations, let Y be a closed subvariety of X and $\mathcal{Y}$ be its Zariski closure in $\mathcal{X}$, that we will consider as an additional domain to be excluded.

We denote by $M_{\overline{K}}$ the places of $\overline{K}$ and $\mathcal{U} = (U_v)_{v \in M_{\overline{K}}}$ a familiy of respective neighbourhoods of $Y(\overline{K}_v)$ in the v-adic topology of $X(\overline{K}_v)$ (Definition III.1.4) which have to be, in some sense, " uniformly large enough around Y " (for the precise meaning of this, see Definition III.4.5). We can see, if Y is a curve, as each neighbourhood U_v as a tube around this curve in X, hence the name *tubular neighbourhood of* Y we give to such a family $\mathcal{U}$ (Remark III.4.5). We then say that a point $P \in X(\overline{K})$ belongs to this tubular neighbourhood if and only if it belongs to a U_v for some $v \in M_{\overline{K}}$. Our theorem can then be stated in the following way.

Theorem (Tubular Runge)**.** *With the above notations, let $\mathcal{E}(L, S_L)(\mathcal{U})$ be the set of points of $(\mathcal{X}\backslash\mathcal{D})(\mathcal{O}_{L,S_L})$ which moreover have the property of being v-adically far away from Y at every place of $M_{\overline{K}}$, that is not belonging to $\mathcal{U}$. Then, for every fixed tubular neighbourhood $\mathcal{U}$ of Y, the union of all the $\mathcal{E}(L, S_L)(\mathcal{U})$ is finite, with (L, S_L) running through the pairs satisfying the tubular Runge condition*

$$m_Y |S_L| < r, \tag{8}$$

where m_Y is the smallest integer such that the intersection of any $(m_Y + 1)$ distinct divisors among $D_1, \cdots, D_r$ is included in Y (and not empty as in general Runge condition).

Notice that if $\mathcal{Y}$ is included in $\mathcal{D}$ (which is the case we have in mind), we know that the $P \in (\mathcal{X}\backslash\mathcal{D})(\mathcal{O}_{L,S_L})$ are not v-adically too close to Y for $v \in M_L \backslash S_L$. The condition exposed then completes the integrality hypotheses at the places of S_L, in particular at every archimedean place.

This result is indeed a concentration result, because it means that for every tubular neighbourhood $\mathcal{U} = (U_v)_{v \in M_{\overline{K}}}$, up to a finite number of points, the $P \in (\mathcal{X}\backslash\mathcal{D})(\mathcal{O}_{L,S_L})$, where (L, S_L) satisfies Runge condition, are *concentrated* near Y, that is belong to the tubular neighbourhood $\mathcal{U}$. The complete statement (more general and precise) of tubular Runge theorem is the Theorem III.8, and its proof is the goal of section III.4.

This theorem is automatic when Y is ample, as well as Northcott property for the height associated to an ample divisor (proposition III.4.6). On the contrary, when Y is empty, it is exacctly Theorem III.7. To apply this hybrid theorem, we then have to make a compromise to have both Y and m_Y as small as possible.

This theorem can also be compared to those of [CLZ09], notably to Autissier's Theorem (for surfaces) and CLZ Theorem (Corvaja, Levin and Zannier) for varieties of dimension at least 3. These last two theorems have a similar flavour to our concentration result, but here are some notable differences. First, they do not need a condition on the size of S_L, nor a condition of intersection as strict as in our theorem, nor an additional hypothesis of exclusion of a tubular neighbourhood. On the other hand, the finiteness results stated depend on S_L (though their exclusion domain Y is absolute) : in particular, it does not say if the reunion of all this finite sets is finite itself, and we do not know what happens in the case of big divisors. Finally, as the foundation of these results is Schmidt's subspace theorem, it seems difficult to make them effective, as opposed to our theorem.

The objective of section III.5 is an application of our tubular Runge theorem to Siegel modular varieties (Theorem III.10), and we here state its more striking case, that is in level $n = 2$ (Theorem III.9). The variety $A_2(2)_{\mathbb{C}}$ is the Siegel modular variety associated to principally polarised abelian surfaces endowed with a symplectic structure of level 2 (Definition-proposition III.5.3), and $A_2(2)^S_{\mathbb{C}}$ is its Satake compactification (Definition III.5.6). Both varieties admit natural normal proper models on $\mathbb{Z}[1/2]$ (Definition-proposition III.5.14).

In this setting, the "boundary" (locus relatively to which we define the integrality) has a good modular interpretation as a locus of degeneracy of abelian surfaces. Indeed, it is made up of the "strong boundary", locus of semiabelian surfaces with nonzero toric part, and of the "weak boundary", locus of nonsimple principally polarised abelian surfaces, whose components make up the effective divisors defining the usual integrality for the theorem.

This situation is analogous to the case of modular curves, where the boundary is the locus of degenerate elliptic curves, that is the cusps, and there is no difference between weak and strong boundary.

Theorem (Runge for products of elliptic curves on $\mathcal{A}_2(2)^S$)**.**

For an open neighbourhood U of the strong boundary $\partial A_2(2)_{\mathbb{C}} := A_2(2)^S_{\mathbb{C}} \backslash A_2(2)_{\mathbb{C}}$ in the usual complex topology, let $\mathcal{E}(U)$ be the set of points $P \in A_2(2)(\overline{\mathbb{Q}})$ associated to triples (A, λ, α_2) (up to isomorphism on some scalar extension), with A an abelian surface, λ a principal polarisation on A and α_2 a symplectic structure of level 2 on A such that :

- *$P \in A_2(2)(L)$ for some number field L.*
- *A has potentially good reduction at every finite place, and for any scalar extension $\sigma : L \to \mathbb{C}$, the point $P_\sigma \in A_2(2)_{\mathbb{C}}$ does not belong to U.*
- *The number s_L of places v of L such that*
  - *v is archimedean*
  - *v divides 2*
  - *The reduction mod v of (A, λ) degenerates (up to scalar extension) to a product of elliptic curves (that is, belongs to the weak boundary)*

satisfies the condition

$$s_L < 10.$$

Then, the set $\mathcal{E}(U)$ is finite for any open neighbourhood U of $\partial A_2(2)_{\mathbb{C}}$.

In this application of our theorem, we used a natural tubular neighbourhood of $\partial\mathcal{A}_2(2)_{\mathbb{Q}}$. Indeed, for every finite place $v \in M_{\overline{\mathbb{Q}}}$ associated to some prime ideal $\mathfrak{P}$, we can consider U_v the set of points $P_v \in A_2(2)^S(\overline{\mathbb{Q}}_v)$ whose reduction modulo $\mathfrak{P}$ belongs to $\partial A_2(2)_{\mathfrak{P}}$. If P_v represents the abelian surface A_v, this means that A_v does not have potentially good reduction modulo $\mathfrak{P}$, by definition of the schemes $\mathcal{A}_2(2)$ and $\mathcal{A}_2(2)^S$ (we can salvage this vision for v dividing 2 even as $\mathcal{A}_2(2)$ is not defined in characteristic 2). There now simply remains to complete this family of open subsets with open subsets at every archimedean place v, and for this a simple means is to fix a unique open neighbourhood U of $\partial A_2(2)_{\mathbb{C}}$, hence the above formulation. The Runge condition $s_L < 10$ comes from the fact that the underlying divisors (making up the weak boundary) are ten, all ample and effective on $A_2(2)^S_{\mathbb{Q}}$ and with empty intersection outside $\partial A_2(2)_{\mathbb{Q}}$ (Definition-proposition III.5.29 (b)), so $m_{\partial A_2(2)} = 1$.

At last, notice that we give in paragraphs III.5.1 and III.5.2 a synthesis of descriptions of good toroidal compactifications of some moduli spaces, as well as of the geometry of their divisors, as a preliminary work for future applications of our tubular Runge theorem to other Siegel modular varieties, and other divisors.

I

Représentations galoisiennes de $\mathbb{Q}$-courbes quadratiques

« La connaissance est un moyen de se nourrir »
– Friedrich Nietzsche

Le problème d'uniformité de Serre

Étant donné un nombre premier p et une courbe elliptique E définie sur un corps de nombres K, soit la représentation

$$\rho_{E,p} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(E[p])$$

obtenue par l'action naturelle de $\mathrm{Gal}(\overline{K}/K)$ sur $E[p]$ le groupe des points de p-torsion de E. Un pas important vers la preuve du grand théorème de Fermat, apporté par Mazur en 1977 [Maz77], consiste à montrer que cette représentation est automatiquement irréductible lorsque $K = \mathbb{Q}$ et E est sans multiplication complexe, pour $p > 163$ (et même $p > 3$ pour une courbe elliptique de Frey associée à une solution théorique de l'équation de Fermat).

Dans un contexte plus général, Serre a montré en 1972 [Ser72] que pour une courbe elliptique E sur K sans multiplication complexe, il existe une borne $C(E, K)$ dépendant de E et K telle que $\rho_{E,p}$ est automatiquement surjective pour $p > C(E, K)$. En fait, nous obtenons dans la section I.8 (théorème I.7), via un théorème d'isogénie, une version explicite du théorème de surjectivité, à savoir que si $h_{\mathcal{F}}(E)$ est la hauteur de Faltings stable de E et p est non ramifié dans K, alors $\rho_{E,p}$ est surjective dès que

$$p > 10^7 [K : \mathbb{Q}]^2 \left(\max\{h_{\mathcal{F}}(E), 985\} + 4 \log[K : \mathbb{Q}]\right)^2 .$$

Le *problème d'uniformité de Serre* consiste alors à savoir si on peut rendre cette surjectivité uniforme en E, c'est-à-dire trouver une borne $C(K)$ dépendant uniquement de K, telle que pour toute courbe elliptique E définie sur K sans multiplication complexe, $\rho_{E,p}$ est surjective pour $p > C(K)$.

Le but du présent travail n'est pas le problème d'uniformité de Serre sur les courbes sur $\mathbb{Q}$ (encore hors de portée des méthodes actuelles, voir la section I.6), mais sur un type légèrement différent de courbes elliptiques : les $\mathbb{Q}$-courbes.

Définition. Une $\mathbb{Q}$-*courbe* est une courbe elliptique E définie sur un corps de nombres K telle que pour tout $\sigma \in \mathrm{Gal}(\overline{K}/\mathbb{Q})$, la conjuguée ${}^{\sigma}E$ est isogène à E (sur $\overline{K}$). Le *degré* de la $\mathbb{Q}$-courbe, noté $d(E)$, est le plus petit multiple commun des degrés des isogénies minimales entre E et ses conjuguées.

• Les $\mathbb{Q}$-courbes généralisent les courbes elliptiques définies sur $\mathbb{Q}$, et on peut associer de telles courbes à des solutions d'équations diophantiennes ternaires (selon [Ell04] ou [BD10] par exemple).

• L'ensemble des $\mathbb{Q}$-courbes est stable par isogénie et contient les courbes elliptiques définies sur $\mathbb{Q}$ mais aussi toutes les courbes elliptiques à multiplication complexe (exemple I.1.1).

• Une $\mathbb{Q}$-courbe est *stricte* si elle n'est pas isogène à une courbe elliptique définie sur $\mathbb{Q}$.

• Pour une $\mathbb{Q}$-courbe E sans multiplication complexe et p premier ne divisant pas $d(E)$, on définit (paragraphe I.1.1) une représentation projective

$$\mathbb{P}\overline{\rho}_{E,p} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}(E[p])$$

analogue à $\rho_{E,p}$ pour les courbes définies sur $\mathbb{Q}$.

• Pour ce qui est de la modularité, Ribet a montré dans [Rib04], avec l'aide de la conjecture de Serre aujourd'hui prouvée ([KW09a] et [KW09b]), que les $\mathbb{Q}$-courbes sont exactement les courbes elliptiques *modulaires* (pour GL_2 sur $\mathbb{Q}$), c'est-à-dire celles qui apparaissent comme quotients de $J_1(N)$ pour un certain N.

Le résultat principal de notre travail sur les $\mathbb{Q}$-courbes et le but de ce chapitre est le théorème suivant.

Théorème I.1 (Surjectivité uniforme pour les $\mathbb{Q}$-courbes strictes)**.**
Soit K un corps quadratique imaginaire de discriminant $-D_K$ et E une $\mathbb{Q}$-courbe stricte définie sur K. Pour tout nombre premier p ne divisant pas $D_K d(E)$ et tel que

$$p > \max(2 \cdot 10^{13}, 50 D_K^{1/4} \log D_K),$$

la représentation $\mathbb{P}\overline{\rho}_{E,p}$ est surjective.

La démarche de la preuve est similaire à celle employé par Bilu et Parent [BP11a] pour la partie « normalisateur de Cartan déployé » du problème d'uniformité sur $\mathbb{Q}$. On fixe un corps quadratique K et pour les assertions suivantes, E parcourt les $\mathbb{Q}$-courbes strictes sans multiplication complexe définies sur K et p ne divise pas $D_K d(E)$. La preuve se décompose alors en quatre grandes étapes.

(***I***) Démontrer par la méthode de Mazur (sections I.2 à I.6) que pour $p > C_0(K)$, si $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas surjective, E a potentiellement bonne réduction en tout idéal de $\mathcal{O}_K$, c'est-à-dire que $j(E) \in \mathcal{O}_K$ (théorèmes I.3, I.4 et I.6). La preuve consiste en un raffinement des résultats de [Ell04], utilisant la structure du groupe des composantes de $J_0(p)_{\overline{\mathbb{F}_p}}$ et le quotient d'Eisenstein (section I.3). Pour le cas « normalisateur de Cartan non déployé » de la méthode, il faut également utiliser une estimation de moyennes de valeurs spéciales de fonctions L (théorème II.1), qui est l'objet du chapitre II.

(***II***) Employer la méthode de Runge (section I.7) pour majorer le j-invariant d'un point P de $X_0(d)(K)$ tel que $j(P) \in \mathcal{O}_K$. On en déduira que pour un tel point,

$$\log |j(P)| \leq C\sqrt{d},$$

pour une certaine constante absolue C (proposition I.7.7 pour l'inégalité précise).

(***III***) Utiliser un théorème d'isogénie à la Masser-Wüstholtz (proposition I.8.1) presque entièrement tiré de [GR14], qui permet de minorer la hauteur du j-invariant d'une courbe elliptique *E sans multiplication complexe* en fonction du degré d'une isogénie de E (théorème I.7). Concrètement, le théorème I.7 (de surjectivité explicite) donnera dans notre situation des bornes de la forme

$$\log |j(E)| \geq C'\sqrt{d(E)p},$$

si $j(E) \in \mathcal{O}_K$ (corollaire I.8.1).

(***IV***) Les trois étapes précédentes nous donnent, si E est une $\mathbb{Q}$-courbe stricte sur un corps K quadratique imaginaire, une borne $C_1(K)$ telle que pour tout nombre premier $p > C_1(K)$, si $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas surjective, $j(E)$ est entier par l'étape (***I***) et comme le noyau C de l'isogénie minimale entre E et sa conjuguée fournit un sous-groupe d'ordre $d(E)$ défini sur K (preuve de la proposition

I.1.17), on peut appliquer l'étape (***II***) avec $(E, C) \in X_0(d(E))(K)$, qui combinée à l'étape (***III***) nous donne des inégalités de la forme

$$C'\sqrt{d(E)p} \leq \log|j(E)| \leq C\sqrt{d(E)},$$

ce qui est impossible pour p assez grand (indépendant de $d(E)$). Les bornes seront explicitées dans la section I.8.

Avant de passer au détail de la preuve, quelques remarques s'imposent sur le théorème I.1.

- Le fait que la $\mathbb{Q}$-courbe E soit stricte implique que le théorème ne dit rien sur les courbes elliptiques définies sur $\mathbb{Q}$, ce qui est naturel puisque le cas « normalisateur de Cartan non déployé » résiste encore pour celles-ci, en particulier à la méthode de Mazur. Un suivi attentif de chacun des sections permettrait à peu de choses près de redémontrer directement les résultats de surjectivité uniforme partielle établis successivement par [Maz77] (« cas Borel ») puis [Mom84] et [BP11a] (« cas normalisateur de Cartan déployé »).

- Le corps quadratique K doit être imaginaire à la fois pour le cas « normalisateur de Cartan non déployé » (section I.6) et la méthode de Runge (section I.7). Pour tout le reste, le cas quadratique réel fonctionne exactement de la même manière. De plus, on a besoin que p soit non ramifié dans K à la fois pour que $\mathbb{P}\overline{\rho}_{E,p}$ ne soit pas inclus dans $\mathrm{PSL}(E[p])$ (paragraphe I.1.2) et pour les estimations analytiques de fonctions L du chapitre II. Il est cependant possible que pour p divisant D_K, les fonctions L en question vérifient d'autres équations fonctionnelles permettant quand même ce type d'estimations. De plus, la dépendance en D_K de la borne vient seulement du terme d'erreur des estimations utilisées, il est donc peut-être possible d'obtenir un résultat avec une dépendance faible en D_K via des techniques différentes.

- Il n'y a aucune dépendance en le degré d de la $\mathbb{Q}$-courbe stricte, à part le fait que p ne doit pas diviser d (la représentation n'est de toute façon pas définie telle quelle si p divise d). Cela permet une certaine marge de manoeuvre sur le résultat, car hormis des cas de petits degrés (Exemple I.1.2), on ne sait pas exhiber de familles infinies de $\mathbb{Q}$-courbes strictes de degré fixé, et la finitude de tels ensembles de $\mathbb{Q}$-courbes ferait du théorème un corollaire immédiat de la surjectivité de Serre.

Notations

Nous regroupons ici les notations les plus utilisées pour des raisons de clarté. On note

p	un nombre premier fixé
E	une courbe elliptique définie sur $K \subset \overline{\mathbb{Q}}$.
${}^{\sigma}E$	la conjuguée galoisienne de E par $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.
$E[p]$	la p-torsion de E, non canoniquement isomorphe à $\left(\mathbb{Z}/p\mathbb{Z}\right)^2$.
$\rho_{E,p}$	la représentation galoisienne de $\mathrm{Gal}(\overline{K}/K)$ sur cette p-torsion.
$\mathbb{P}\overline{\rho}_{E,p}$	la représentation galoisienne projective de la $\mathbb{Q}$-courbe E.

I.1 Rappels sur les $\mathbb{Q}$-courbes et le problème de surjectivité

I.1.1 Représentations projectives de $\mathbb{Q}$-courbes

Pour commencer, donnons en détail la définition de $\mathbb{Q}$-courbe et la terminologie associée.

Définition I.1.1 ($\mathbb{Q}$-courbes)**.** Soit K un corps de nombres.

Une $\mathbb{Q}$-*courbe* est une courbe elliptique E définie sur K telle que pour tout $\sigma \in \mathrm{Gal}(\overline{K}/\mathbb{Q})$, il existe une isogénie $\mu_\sigma : {}^{\sigma}E \to E$ (définie sur $\overline{K}$). Elle est *quadratique* si le corps K est quadratique.

Son *degré*, noté $d(E)$, est le ppcm des degrés des isogénies minimales entre E et ses conjuguées galoisiennes.

Une $\mathbb{Q}$-courbe est *stricte* si elle n'est pas isogène à une courbe elliptique définie sur $\mathbb{Q}$.

Remarque I.1.1. Quitte à étendre le corps de base K, on peut supposer que celui-ci est une extension *galoisienne* de $\mathbb{Q}$ de sorte que l'action de $\mathrm{Gal}(\overline{K}/\mathbb{Q})$ sur les conjuguées de E se factorise par $\mathrm{Gal}(K/\mathbb{Q})$. Alors, une courbe elliptique E définie sur le corps de nombres galoisien K est une $\mathbb{Q}$-courbe si et seulement si pour tout $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, il existe une isogénie entre ${}^{\sigma}E$ et E.

Exemple I.1.1 (Courbes CM)**.** Les courbes elliptiques sur $\mathbb{Q}$ sont des $\mathbb{Q}$-courbes. Il en est de même pour les courbes à multiplication complexe : en effet, si $\mathrm{End}\, E = A \neq \mathbb{Z}$, pour tout automorphisme σ de $\overline{K}$, E^{σ} est également à multiplication complexe par A ([Sil94], Proposition II.2.1 (a)), or on a une action (simplement) transitive du groupe des classes d'idéaux de A sur l'ensemble des courbes elliptiques à multiplication complexe par A à isomorphisme près, et cette action vient avec des isogénies entre toutes ces courbes ([Sil94], Proposition II.1.2 (b)). Ainsi, E est une $\mathbb{Q}$-courbe.

Pour N sans facteur carré, et $X_0^*(N)$ le quotient de la courbe modulaire $X_0(N)$ par le groupe de ses involutions d'Atkin-Lehner, un point non cuspidal de $X_0^*(N)(\mathbb{Q})$ se relève nécessairement en une $\mathbb{Q}$-courbe (unique à isogénie près). En fait, Elkies a démontré qu'on obtient ainsi toutes les classes d'isogénie de $\mathbb{Q}$-courbes non CM, avec le résultat suivant [Elk04].

Théorème I.2 (Réduction d'Elkies)**.**
Soit E une $\mathbb{Q}$-courbe sans multiplication complexe définie sur un corps de nombres K.
Alors, il existe N sans facteur carré divisant $d(E)$ et E' une $\mathbb{Q}$-courbe définie sur un sous-corps de K et associée à un point de $X_0^*(N)(\mathbb{Q})$, telle que E' est isogène à E sur $\overline{\mathbb{Q}}$, via une isogénie de degré divisant $d(E)$.

Il existe bien une infinité de $\mathbb{Q}$-courbes strictes et sans multiplication complexe, comme les familles paramétrées suivantes en donnent. Remarquons en préambule qu'une $\mathbb{Q}$-courbe E non définie sur $\mathbb{Q}$ et sans multiplication complexe dont le degré est sans facteur carré est nécessairement stricte, sinon elle admettrait une isogénie $E \to E$ de degré non carré.

Exemple I.1.2 (Familles paramétrées de $\mathbb{Q}$-courbes)**.**
• Une première classe de familles de $\mathbb{Q}$-courbes quadratiques paramétrées ([Has97], Théorèmes 2.2 et 2.4) est donnée par les $X_0(p)$ de genre 0 avec p premier, c'est-à-dire pour $p = 2, 3, 5, 7, 13$. Lorsque $p = 2, 3$ ou 7, on sait fabriquer une famille paramétrée de $\mathbb{Q}$-courbes définies sur $\mathbb{Q}(\sqrt{d})$ pour tout $d \neq 0, 1$ sans facteur carré, par exemple pour $p = 2$ on a pour tout $t \in \mathbb{Q}$ la $\mathbb{Q}$-courbe définie sur $\mathbb{Q}(\sqrt{d})$ par l'équation

$$E_d(t) : \quad y^2 = x^3 + 6(3t\sqrt{d} - 5)x + 8(-9t\sqrt{d} + 7), \qquad j(E_d(t)) = 2^6 \frac{(5 - 3t\sqrt{d})^3}{(1 - t\sqrt{d})(1 + t\sqrt{d})^2}.$$

Pour $p = 5$ ou 13 apparaissent des conditions de congruence sur d pour l'existence de $\mathbb{Q}$-courbes non triviales ([Has97], Proposition 2.3) mais sous réserve qu'elles soient satisfaites, il existe également sur $\mathbb{Q}(\sqrt{d})$ une famille de $\mathbb{Q}$-courbes paramétrée par $\mathbb{Q}$.

Remarquons que les degrés des $\mathbb{Q}$-courbes sont ici des nombres premiers donc à part pour les courbes CM, elles sont nécessairement strictes. Les courbes CM sont pour leur part en nombre fini (à d fixé, pour chaque famille de paramètres), ainsi pour $p = 2, 3, 5, 7, 13$, on a bien des familles infinies de $\mathbb{Q}$-courbes strictes de degré p sur des corps quadratiques fixés. De plus, la méthode employée dans [Has97] démontre que toutes les $\mathbb{Q}$-courbes quadratiques de ces degrés s'obtiennent à isogénie près exactement de cette manière.

• Plus généralement, lorsque la courbe $X_0^*(N)$, quotient de $X_0(N)$ par son groupe d'Atkin-Lehner (N supposé sans facteur carré) est de genre 0 ou 1, on peut également paramétrer d'après [GL98] toutes les $\mathbb{Q}$-courbes (via leurs j-invariants comme racines de certains polynômes à coefficients rationnels), mais alors le corps de définition de la $\mathbb{Q}$-courbe (c'est-à-dire son j-invariant) dépend du paramètre hors des cas précédents, c'est-à-dire qu'on ne sait pas a priori s'il existe, à corps quadratique fixé, une infinité de $\mathbb{Q}$-courbes de degré fixé définies sur ce corps. Une conjecture d'Elkies [Elk04] (toujours non résolue même pour N premier) affirme en fait que pour tout corps de nombres K fixé, la courbe $X_0^*(N)$ n'a pas de point K-rationnel ni cuspidal ni CM, pour N assez grand.

Nous allons dorénavant supposer les $\mathbb{Q}$-courbes étudiées *sans multiplication complexe*.

Définition I.1.2 (Représentations projectives associées à une $\mathbb{Q}$-courbe)**.**
Soient $K \subset \overline{\mathbb{Q}}$ un corps de nombres et E une $\mathbb{Q}$-courbe *sans multiplication complexe* définie sur K. Pour tout nombre premier p ne divisant pas $d(E)$, la formule

$$\sigma \cdot D := \mu_\sigma({}^\sigma D),$$

avec D parcourant les $\mathbb{F}_p$-droites de $E[p]$, σ parcourant $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ et $\mu_\sigma : {}^\sigma E \to E$ une isogénie de degré premier à p, définit une représentation de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ dans $\mathbb{P}E[p]$ notée $\mathbb{P}\overline{\rho}_{E,p}$ qui ne dépend pas du choix des μ_σ et qui, restreinte à $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$, est la projectivisation du morphisme $\rho_{E,p}$ de $\mathrm{Gal}(\overline{K}/K)$ dans $\mathrm{GL}(E[p])$.

Démonstration. Soit D une $\mathbb{F}_p$-droite de $E[p]$. Pour tout $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, ${}^\sigma D$ est une $\mathbb{F}_p$-droite de ${}^\sigma E$ et l'isogénie μ_σ est injective sur la p-torsion de ${}^\sigma E$ donc $\sigma \cdot D$ est bien une $\mathbb{F}_p$-droite de $E[p]$. L'application $\mathbb{P}\overline{\rho}_{E,p}$ est donc bien définie. Elle ne dépend pas du choix des isogénies : en effet, comme E et ${}^\sigma E$ sont sans multiplication complexe, $\mathrm{Hom}({}^\sigma E, E)$ est un $\mathbb{Z}$-module libre de rang 1, en particulier toutes les isogénies entre ${}^\sigma E$ et E de degré premier à p agissent de la même manière sur les $\mathbb{F}_p$-droites de ${}^\sigma E[p]$.

Ensuite, pour tous $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, comme les isogénies $\mu_{\sigma\tau}$ et $\mu_\sigma \circ {}^\sigma\mu_\tau$ de ${}^{\sigma\tau}E$ vers E sont de degré premier à p, on a

$$(\sigma\tau)\cdot D = \mu_{\sigma\tau}({}^{\sigma\tau}D) = \mu_\sigma \circ {}^\sigma\mu_\tau({}^{\sigma\tau}D) = \mu_\sigma({}^\sigma(\mu_\tau({}^\tau D))) = \sigma\cdot(\tau\cdot D)$$

donc $\mathbb{P}\overline{\rho}_{E,p}$ est bien un morphisme de groupes. □

Le théorème I.2 a un corollaire immédiat très utile pour notre problème de surjectivité.

Corollaire I.1.1. *Avec les notations de la proposition précédente, pour p premier ne divisant pas $d(E)$, il suffit de montrer que $\mathbb{P}\overline{\rho}_{E',p}$ est surjective pour que $\mathbb{P}\overline{\rho}_{E,p}$ le soit. Il suffit donc de démontrer le théorème I.1 pour les $\mathbb{Q}$-courbes E telles que $d(E)$ est sans facteur carré.*

Démonstration. Soit $\mu : E \to E'$ l'isogénie obtenue grâce au théorème I.2. Alors, pour tout $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ et toute $\mathbb{F}_p$-droite de $E[p]$, pour des isogénies $\mu_\sigma : {}^\sigma E \to E$ et $\mu'_\sigma : {}^\sigma E' \to E'$ de degré premier à p, on a

$$\sigma(\mu(D)) = \mu'_\sigma({}^\sigma(\mu(D))) = \mu'_\sigma({}^\sigma\mu({}^\sigma D)) = \mu'_\sigma \circ {}^\sigma\mu({}^\sigma D) = \mu\circ\mu_\sigma({}^\sigma D) = \mu(\sigma(D))$$

car les isogénies $\mu\circ\mu_\sigma$ et $\mu'_\sigma\circ{}^\sigma\mu$, de degré premier à p, agissent de même sur $\mathbb{P}E[p]$. Ceci implique que la bijection $D \mapsto \mu(D)$ est équivariante pour les actions de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ sur $\mathbb{P}E(p]$ et $\mathbb{P}E'[p]$, donc les images des représentations associées sont simultanément surjectives. □

I.1.2 Découpage du problème de surjectivité

Le problème de surjectivité de $\mathbb{P}\overline{\rho}_{E,p}$ passe par la recherche des sous-groupes maximaux de $\mathrm{PGL}_2(\mathbb{F}_p)$, car $E[p] \cong {\mathbb{F}_p}^2$ après choix de base de la p-torsion ([Sil09], Corollaire III.6.4). Pour prouver que $\mathbb{P}\overline{\rho}_{E,p}$ est surjective, il suffira de montrer que son image n'est pas incluse dans un de ces sous-groupes, que nous allons définir ci-dessous.

Définition I.1.3 (Sous-groupes remarquables de $\mathrm{GL}_2(\mathbb{F}_p)$)**.**
Soit V un $\mathbb{F}_p$-espace vectoriel de dimension 2 et $\pi : \mathrm{GL}(V) \to \mathrm{PGL}(V)$ la projection canonique.

- Un *sous-groupe de Borel* de $\mathrm{GL}(V)$ est le stabilisateur d'une droite D de V pour l'action naturelle de $\mathrm{GL}(V)$. Après un choix de base adaptée de V, un sous-groupe de Borel est le groupe des matrices triangulaires supérieures de $\mathrm{GL}_2(\mathbb{F}_p)$.

- Un *sous-groupe de Cartan déployé* est le stabilisateur d'un couple de droites distinctes (D_1, D_2) de V pour l'action produit de $\mathrm{GL}(V)$. Après un choix de base adaptée de V, c'est le groupe des matrices diagonales de $\mathrm{GL}_2(\mathbb{F}_p)$.

- Un *sous-groupe de Cartan non déployé* est une copie du groupe $\mathbb{F}_{p^2}^*$ dans $\mathrm{GL}(V)$.

On appelle *sous-groupe de Borel* (resp. *Cartan déployé*, *Cartan non déployé*) de $\mathrm{PGL}(V)$ l'image par π d'un sous-groupe de Borel (resp. Cartan déployé, Cartan non déployé) de $GL(V)$.

Remarque I.1.2. Si une sous-algèbre de $\mathrm{End}(V)$ est un corps à p^2 éléments, l'ensemble de ses inversibles forme un sous-groupe de Cartan non déployé de $\mathrm{GL}(V)$. Réciproquement, si C est un sous-groupe de Cartan non déployé, comme il est commutatif et que $X^{p^2-1}-1$ est scindé à racines simples dans $\mathbb{F}_p$, tous ses éléments sont codiagonalisables dans une base de $V \otimes_{\mathbb{F}_p} {\mathbb{F}_p}^2$ formée de vecteurs conjugués, et on en déduit que $C \cup \{0\}$ est un corps à p^2 éléments dans $\mathrm{End}(V)$. De plus, un élément de $\mathrm{GL}(V)$ appartient à C si et seulement s'il commute avec les éléments de C, c'est-à-dire si et seulement s'il agit linéairement sur V pour la structure de $\mathbb{F}_{p^2}$-espace vectoriel de V définie par C. En effet, si g commute avec les éléments de C, il est automatiquement diagonalisable dans une base de $V \otimes_{\mathbb{F}_p} {\mathbb{F}_p}^2$ adaptée à C (donc formée de vecteur conjugués), donc déterminé par ses valeurs propres, or C couvre les p^2-1 possibilités donc $g \in C$.

En fait, on peut également voir un sous-groupe de Cartan non déployé de $\mathrm{GL}(V)$ comme le groupe des matrices de $\mathrm{GL}(V)$ diagonales dans une base adaptée à (D, D') où D et D' sont des droites de $V \otimes_{\mathbb{F}_p} \mathbb{F}_{p^2}$ distinctes conjuguées.

Lemme I.1.4 (Normalisateurs des sous-groupes remarquables)**.**

Soit V un $\mathbb{F}_p$-espace vectoriel de dimension 2.

- Le normalisateur d'un sous-groupe de Borel de $\mathrm{GL}(V)$ *(resp.* $\mathrm{PGL}(V)$*) est réduit à lui-même.*

- Le normalisateur N du sous-groupe de Cartan déployé C de $\mathrm{GL}(V)$ *(resp.* $\mathrm{PGL}(V)$*) associé au couple de droites distinctes (D_1, D_2) de V est le stabilisateur de la paire de droites distinctes $\{D_1, D_2\}$. En particulier, $[N : C] = 2$.*

- Le normalisateur N du sous-groupe de Cartan non déployé C de $\mathrm{GL}(V)$ *(resp.* $\mathrm{PGL}(V)$*) est l'ensemble des éléments de* $\mathrm{GL}(V)$ *qui agissent $\mathbb{F}_{p^2}$-linéairement ou semi-linéairement par rapport à la structure de $\mathbb{F}_{p^2}$-espace vectoriel de V donnée par C. En particulier, $[N : C] = 2$ (et de même pour* $\mathrm{PGL}(V)$*).*

Enfin, l'intersection de deux sous-groupes de Cartan distincts de $\mathrm{GL}(V)$ *est réduite aux homothéties.*

Démonstration. Si pour toute droite D de V on note B_D le sous-groupe de Borel associé à D, on voit immédiatement que pour tout $g \in \mathrm{GL}(V)$, $gB_Dg{-}1 = B_{gD}$. En particulier, $gB_Dg^{-1} = B_D$ si et seulement si g fixe D, c'est-à-dire $g \in B_D$.

Pour une paire de droites distinctes $\{D_1, D_2\}$ de V, si on note $C = C_{\{D_1,D_2\}}$ le sous-groupe de Cartan déployé associé à cette paire de droites, on a $gC_{\{D_1,D_2\}}g^{-1} = C_{\{gD_1,gD_2\}}$, en particulier g normalise C si et seulement si $\{gD_1, gD_2\} = \{D_1, D_2\}$. Dans une base adaptée à (D_1, D_2), C est constitué des matrices diagonales et son normalisateur N des matrices diagonales ou antidiagonales, donc $[N : C] = 2$.

Pour un sous-groupe de Cartan non déployé C de $\mathrm{GL}(V)$, le raisonnement ci-dessus se reproduit exactement lorsqu'on utilise la remarque I.1.2.

Enfin, pour ce qui est de l'intersection de deux sous-groupes de Cartan, si un élément de $\mathrm{GL}(V)$ appartient à un sous-groupe de Cartan, il possède comme droites propres les droites associées à ce sous-groupe, or s'il a au moins trois droites propres, c'est une homothétie.

□

La proposition suivante provient de ([Ser72], § 2.4 à 2.6), sa preuve y est esquissée mais nous en reproduisons ici une démonstration complète pour le confort du lecteur, basée sur la proposition 15 de [Ser72] pour le (a) et la preuve du théorème 6.17 de [Suz82] pour le (b).

Proposition I.1.5 (Sous-groupes maximaux de $\mathrm{GL}_2(\mathbb{F}_p)$)**.**

Soit H un sous-groupe de $\mathrm{PGL}_2(\mathbb{F}_p)$*.*

(a) Si H est d'ordre divisible par p, alors ou bien H contient $\mathrm{PSL}_2(\mathbb{F}_p)$*, ou bien H est contenu dans un sous-groupe de Borel de* $\mathrm{PGL}_2(\mathbb{F}_p)$*.*

(b) Si H est d'ordre premier à p, trois cas sont possibles :

- *Si H est cyclique, il est contenu dans un sous-groupe de Cartan de* $\mathrm{PGL}_2(\mathbb{F}_p)$*, unique si H est non trivial.*
- *Si H est diédral, il normalise un sous-groupe de Cartan de* $\mathrm{PGL}_2(\mathbb{F}_p)$*, il est donc contenu dans le normalisateur de ce sous-groupe de Cartan.*

– *Si H n'est cyclique ni diédral, il est isomorphe à $\mathfrak{A}_4$, $\mathfrak{S}_4$ ou $\mathfrak{A}_5$ (on dit alors qu'il est exceptionnel).*

Démonstration. On note $\pi : \mathrm{GL}_2(\mathbb{F}_p) \to \mathrm{PGL}_2(\mathbb{F}_p)$ la projection canonique, $G = \pi^{-1}(H)$ et Z le centre de $\mathrm{GL}_2(\mathbb{F}_p)$. On suppose que H est non trivial, c'est-à-dire que $G \neq Z$.

(a) Comme la seule racine p-ième de l'unité dans $\overline{\mathbb{F}_p}$ est 1, toute matrice g de $\mathrm{GL}_2(\mathbb{F}_p)$ d'ordre exactement p est semblable à la matrice $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, en particulier fixe une unique droite notée D_g de ${\mathbb{F}_p}^2$ et pour tout $g' \in G$, $D_{g'gg'^{-1}} = g' \cdot D_g$. Ainsi, soit tous les éléments de G fixent D_g et alors H est inclus dans le sous-groupe de Borel fixant D_g, soit il existe deux éléments g et g' de G d'ordre p qui chacun fixent une droite différente, et alors dans une base adaptée à ces deux droites, ils sont de la forme $\begin{pmatrix}1 & a\\ 0 & 1\end{pmatrix}$ et $\begin{pmatrix}1 & 0\\ b & 1\end{pmatrix}$ avec $ab \neq 0$. Ces matrices engendrent $\mathrm{SL}_2(\mathbb{F}_p)$, donc G contient $\mathrm{SL}_2(\mathbb{F}_p)$ et H contient $\mathrm{PSL}_2(\mathbb{F}_p)$.

(b) Commençons par calculer, pour tout $g \in \mathrm{GL}_2(\mathbb{F}_p)$, le centralisateur de g dans $\mathrm{GL}_2(\mathbb{F}_p)$:

- Si g est une homothétie, c'est tout $\mathrm{GL}_2(\mathbb{F}_p)$.

- Si g a deux valeurs propres distinctes λ et μ dans $\mathbb{F}_p$, c'est le groupe des matrices qui dans une base propre de g sont de la forme $\begin{pmatrix}* & 0\\ 0 & *\end{pmatrix}$, c'est-à-dire un sous-groupe de Cartan déployé de $\mathrm{GL}_2(\mathbb{F}_p)$.

- Si g a deux valeurs propres distinctes λ et μ dans $\mathbb{F}_{p^2}\backslash\mathbb{F}_p$, λ et μ sont conjuguées et le centralisateur de g est alors constitué de toutes les matrices de $\mathrm{GL}_2(\mathbb{F}_p)$ qui dans une base propre de $\mathbb{F}_{p^2}$ associée à g sont de la forme $\begin{pmatrix}\lambda' & 0\\ 0 & \lambda'^p\end{pmatrix}$. C'est donc exactement le sous-groupe de Cartan non déployé de $\mathrm{GL}_2(\mathbb{F}_p)$ associé à ces droites propres.

- Si g n'est pas diagonalisable dans $\overline{\mathbb{F}_p}$, elle est semblable dans $\mathbb{F}_p$ à une matrice de la forme $\begin{pmatrix}\lambda & \mu\\ 0 & \lambda\end{pmatrix}$ avec $\mu \neq 0$, et son centralisateur (par un calcul immédiat) est alors constitué des matrices qui dans cette base de trigonalisation de g sont de la forme $\begin{pmatrix}\lambda' & \mu'\\ 0 & \lambda'\end{pmatrix}$, il est donc de cardinal $(p-1)p$.

Comme H est d'ordre premier à p, G l'est aussi et le quatrième cas ne peut pas se produire pour $g \in G$: en effet, il appartient à $C_G(g)$, or s'il n'est pas diagonalisable dans $\overline{\mathbb{F}_p}$, celui-ci est réduit au centre de $\mathrm{GL}_2(\mathbb{F}_p)$ d'après le calcul précédent, car p ne divise pas l'ordre de G. Tout élément g de G non scalaire admet donc pour centralisateur $C_G(g)$ le sous-groupe d'un groupe de Cartan, en particulier celui-ci est abélien. De plus, le normalisateur $N_G(g)$ de chaque centralisateur vérifie $[N_G(g) : C_G(g)] \leq 2$: en effet, d'après le lemme I.1.4, pour $g \in G$ non scalaire, $C_G(g)$ est contenu dans un unique sous-groupe de Cartan de $\mathrm{GL}(V)$ donc le normaliser équivaut à normaliser ce sous-groupe de Cartan, ainsi $N_G(g)$ est l'intersection du normalisateur de ce sous-groupe de Cartan avec G, donc $[N_G(g) : C_G(g)] \leq 2$.

Les centralisateurs $C_G(g)$ avec g non scalaire constituent tous les sous-groupes abéliens maximaux de G car si $g \in A$ avec A abélien, $A \subset C_G(g)$. De plus, l'intersection de deux sous-groupes abéliens maximaux distincts A et B de G est Z, car A et B sont contenus dans le centralisateur de n'importe quel élément de leur intersection, qui doit donc être réduite à Z par maximalité.

Notons $\mathcal{A}$ l'ensemble des sous-groupes abéliens maximaux de G, et $C_1, \cdots, C_r, \cdots, C_{r+s}$ les classes de conjugaison de ces sous-groupes abéliens (c'est-à-dire que A et A' sont dans la même classe si $A' = gAg^{-1}$ pour un certain $g \in G$), avec A_i un représentant de C_i pour tout i, de sorte que $N_G(A_i) = A_i$ si $1 \leq i \leq r$ et $[N_G(A_i) : A_i] = 2$ si $r+1 \leq i \leq r+s$. On note $n = |H| = |G|/(p-1)$, et $n_i = |A_i|/(p-1)$ pour tout i.

Par définition du normalisateur, le nombre d'éléments de C_i est exactement $|G|/N_G(A_i)$, et chacun des éléments de $G\backslash Z$ appartenant à un seul sous-groupe abélien maximal (qui est son centralisateur), on obtient l'égalité

$$(p-1)n = (p-1) + \sum_{i=1}^{r} \frac{(p-1)(n_i-1)n(p-1)}{(p-1)n_i} + \sum_{i=r+1}^{r+s} \frac{(p-1)(n_i-1)n(p-1)}{2(p-1)n_i}$$

qui se simplifie en

$$1 = \frac{1}{n} + \sum_{i=1}^{r} \frac{n_i - 1}{n_i} + \sum_{i=r+1}^{r+s} \frac{n_i - 1}{2n_i} \tag{I.1}$$

Or, chacun des n_i vaut au moins 2 car $G \neq Z$, donc on obtient l'inégalité

$$\frac{1}{n} + \frac{r}{2} + \frac{s}{4} \leq 1$$

ce qui nous donne six possibilités pour r et s, décrites dans le tableau suivant

	I	II	III	IV	V	VI
r	1	1	0	0	0	0
s	0	0	1	2	3	4

Nous allons réécrire l'égalité (I.1) dans chacun de ces cas et en déduire les différentes possibilités pour H.

Cas I : on a $1 = 1/n + (n_1 - 1)/n_1$ d'où $n_1 = n$ et G est un sous-groupe abélien, en particulier inclus dans un sous-groupe de Cartan (déployé ou non) de $\mathrm{GL}_2(\mathbb{F}_p)$ vu la structure des centralisateurs, et H est alors cyclique.

Cas II : on a $1 = 1/n + (n_1 - 1)/n_1 + (n_2 - 1)/2n_2$ d'où $1/2 + 1/n = 1/n_1 + 1/2n_2$ ce qui donne $n_1 = 2$ et $2n_2 = n$ ou bien $n_1 = 3, n_2 = 2$ et $n = 12$. Dans le premier cas, on obtient $G = N_G(A_2)$ donc H est diédral. Dans le second, il y a exactement quatre conjugués de A_1 dans G et G agit par conjugaison en permutant ceux-ci, or le stabilisateur de chacun de ces groupes est réduit à lui-même et leur intersection réduite au centre donc $H = G/Z$ s'injecte dans $\mathfrak{S}_4$, d'image de cardinal 12 donc H est isomorphe à $\mathfrak{A}_4$.

Cas III, IV et V : on obtient pour chacun de ceux-ci que $n = 1$, contradiction.

Cas VI : on a $1 = 1/n + (n_1 - 1)/2n_1 + (n_2 - 1)/2n_2 + (n_3 - 1)/2n_3$ donc

$$\frac{1}{2n_1} + \frac{1}{2n_2} + \frac{1}{2n_3} = \frac{1}{n} + \frac{1}{2}.$$

Les (n_1, n_2, n_3, n) possibles sont, si on impose $n_1 \leq n_2 \leq n_3$, $(2, 2, n_3, 2n_3)$, $(2, 3, 3, 12)$, $(2, 3, 4, 24)$ et $(2, 3, 5, 60)$. Pour $(2, 2, n_3, 2n_3)$, on a $G = N_G(A_3)$ donc G est encore inclus dans le normalisateur d'un sous-groupe de Cartan, ainsi H est diédral.

Pour $(2, 3, 3, 12)$, les groupes $\pi(A_2)$ et $\pi(A_3)$ sont des 3-Sylow de H donc doivent être conjugués par les théorèmes de Sylow, contradiction donc ce cas ne se produit pas. Pour $(2, 3, 4, 24)$, C_2 est de cardinal $24/6 = 4$ donc l'action par conjugaison sur les éléments de C_2 induit un morphisme de H dans $\mathfrak{S}_4$, mais le noyau de ce morphisme est l'intersection de tous les $\pi(N_G(gA_2g^{-1}))$ qui est le groupe trivial, et par argument de cardinal H est donc isomorphe à $\mathfrak{S}_4$.

Enfin, pour $(2, 3, 5, 60)$, il y a exactement $60/4 = 15$ groupes d'indice 30 de G (ce sont les conjugués de A_1) et chacun a pour normalisateur un groupe d'indice 15 de G, qui n'est pas abélien par maximalité de A_1. Or, leur projection est un 2-Sylow de H de cardinal 4, qui ne peut pas être cyclique car sinon chaque normalisateur serait abélien (son quotient par le centre étant cyclique). Les 2-Sylow S de H sont donc isomorphes à $(\mathbb{Z}/2\mathbb{Z})^2$, ainsi chacun d'entre eux contient exactement 3 sous-groupes d'indice 2, et chacun d'eux est un $\pi(A)$ avec $N_G(A) = \pi^{-1}(S)$. On en déduit qu'il y a exactement $15/3 = 5$ 2-Sylow dans H et ils sont tous conjugués par les théorèmes de Sylow, d'où un morphisme de H dans $\mathfrak{S}_5$ donné par l'action par conjugaison, et ce morphisme est injectif car chaque $N_G(A)$ ne normalise que trois sous-groupes d'indices 30 comme on vient de le montrer donc H s'identifie à un sous-groupe d'indice 2 de $\mathfrak{S}_5$, qui ne peut être que $\mathfrak{A}_5$.

□

La proposition suivante permet de calculer le déterminant de la représentation projective associé à une $\mathbb{Q}$-courbe, et donc son image admissible maximale.

Proposition I.1.6. *Soit E une courbe elliptique définie sur un corps de nombres K.*

Le déterminant de la représentation galoisienne $\rho_{E,p} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(E[p])$ est égal au caractère cyclotomique $\chi : \mathrm{Gal}(\overline{K}/K) \to \mathbb{F}_p^*$, c'est-à-dire que pour tout $\sigma \in \mathrm{Gal}(\overline{K}/K)$ et toute racine p-ième de l'unité ζ,

$$\zeta^{\det(\rho_{E,p}(\sigma))} = \sigma(\zeta).$$

Corollaire I.1.2. *Pour K un corps de nombres quadratique et E une $\mathbb{Q}$-courbe définie sur K, l'image de $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas incluse dans $\mathrm{PSL}(V)$ si p est non ramifié dans K.*

Démonstration.

Notons μ_p le groupe des racines p-ièmes de l'unité dans $\overline{K}$ et $e_p : E[p] \times E[p] \longrightarrow \mu_p$ l'accouplement de Weil ([Sil09], Chapitre III, Proposition 8.1). Soit $\zeta \in \mu_p$. L'accouplement étant non dégénéré, il existe une $\mathbb{F}_p$-base (P, Q) de E_p telle que $e_p(P, Q) = \zeta$. Soit $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ et $\rho(\sigma)$ la matrice de $P \mapsto {}^{\sigma}P$ vue dans la base (P, Q), c'est-à-dire que

$$({}^{\sigma}P, {}^{\sigma}Q) = (P, Q).\rho(\sigma)$$

or l'accouplement de Weil est compatible avec l'action du groupe de Galois, ainsi

$$e_p({}^{\sigma}P, {}^{\sigma}Q) = \sigma(e_p(P, Q))$$

donc si $\rho(\sigma) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_p)$, on a

$$\zeta^{ad-bc} = \sigma(\zeta),$$

ce qu'on voulait démontrer.

Pour le corollaire, ce résultat démontre que si p n'est pas ramifié dans K, le déterminant de $\rho_{E,p}$ est surjectif dans ${\mathbb{F}_p}^*$, en particulier pour une $\mathbb{Q}$-courbe, l'image de la restriction de $\mathbb{P}\overline{\rho}_{E,p}$ à $\mathrm{Gal}(\overline{K}/K)$ n'est pas incluse dans $\mathrm{PSL}(V)$. □

Nous allons dès maintenant écarter une deuxième possibilité de sous-groupe maximal (donnée par la proposition I.1.5), celle des sous-groupes exceptionnels.

I.1.3 Le cas des sous-groupes exceptionnels

Ce paragraphe reprend les résultats de [Ser72] pour le confort du lecteur (pour plus de détails, lire la section 1 de cet article). Il ne s'applique pas uniquement aux $\mathbb{Q}$-courbes, donc ici seulement, on note :

K un corps valué complet pour la valuation discrète normalisée v, de caractéristique 0.

$\mathcal{O}$ son anneau des entiers, d'idéal maximal $\mathfrak{m}$.

k le corps résiduel de $\mathcal{O}$, supposé de caractéristique $p > 0$.

$e = v(p)$ l'indice de ramification absolu de K.

$\overline{K}$ une clôture algébrique de K fixée, d'anneau des entiers $\overline{\mathcal{O}}$, muni de la valuation v prolongée à $\overline{K}$, renotée v.

K_{nr} la plus grande sous-extension non ramifiée de K dans $\overline{K}$, d'anneau des entiers $\mathcal{O}_{nr}$.

K_t la plus grande sous-extension modérément ramifiée de K dans $\overline{K}$, d'anneau des entiers $\mathcal{O}_t$.

Le corps résiduel de $\overline{K}$ est une clôture séparable de K, notée k_s. Via la réduction modulo leurs idéaux maximaux, $\mathcal{O}_{nr}$ et $\mathcal{O}_t$ se surjectent dans k_s.

On a

$$K \subset K_{nr} \subset K_t \subset \overline{K},$$

et chacune de ces extensions de K est galoisienne. On note

$$G := \mathrm{Gal}(\overline{K}/K), \qquad I := \mathrm{Gal}(\overline{K}/K_{nr}), \qquad I_p := \mathrm{Gal}(\overline{K}/K_t).$$

Le groupe I est le groupe d'inertie de G, c'est-à-dire le groupe des automorphismes de $\overline{K}$ qui se réduisent en l'identité sur k_s. Via la réduction en automorphismes de k_s, $\mathrm{Gal}(K_{nr}/K) = G/I$ s'identifie à $\mathrm{Gal}(k_s/k)$. Le groupe I_p est le plus grand pro-p-groupe inclus dans I, et on note

$$I_t := I/I_p = \mathrm{Gal}(K_t/K_{nr})$$

le *groupe d'inertie modérée* de G. C'est un groupe profini commutatif d'ordre premier à p, dont l'étude va nous donner des informations sur les représentations galoisiennes associées à des courbes elliptiques.

Définition I.1.7 (Caractère θ_d)**.**
Pour $d \geq 1$ premier à p, notons μ_d le groupe des racines d-ièmes de l'unité dans $\overline{K}$. Si x est une uniformisante de K_{nr}, l'extension $K_d = K_{nr}(x^{1/d})/K_{nr}$ est modérément totalement ramifiée et galoisienne de groupe de Galois μ_d, et on a l'homomorphisme continu canonique

$$\theta_d : \left| \begin{array}{ccc} I_t & \longrightarrow & \mu_d \\ \sigma & \longmapsto & \frac{\sigma(x^{1/d})}{x^{1/d}} \end{array} \right.$$

qui ne dépend pas du choix de l'uniformisante x ni de sa racine d-ième car $\mu_d \subset K_{nr}$.

Lorsque d est premier à p, le groupe μ_d s'injecte dans k_s par la réduction, et il est alors inclus dans l'unique copie de $\mathbb{F}_q$ dans k_s, lorsque d divise $q-1$. On renote θ_d le caractère de I_t dans k_s^* ainsi obtenu. Ceux-ci permettent d'obtenir tous les caractères continus, comme l'énonce la proposition suivante.

Proposition I.1.8. *Pour tout caractère continu $\theta : I_t \to k_s^*$, il existe $a \in \mathbb{Z}$, et $d \geq 1$ premier à p tels que $\theta = \theta_d^a$. De plus, pour de tels a et d, la classe α de a/d dans $\mathbb{Q}/\mathbb{Z}$ est uniquement déterminée : on dira que l'invariant de θ est α.*

Démonstration. Tout d'abord, un caractère continu de I_t dans k_s^* (muni de la topologie discrète) est à image compacte (car I_t est compact) et discrète, donc finie et dans un certain $\mu_d \subset k_s^*$. Le groupe I_t est la limite projective des $\mathrm{Gal}(K_{nr}(x^{1/n})/K_{nr} = \mu_n$, et on voit immédiatement que l'application qui à $a \in \mathbb{Z}$ associe θ_d^a induit un isomorphisme entre $\mathbb{Z}/d\mathbb{Z}$ et le groupe des caractères de I_t dans μ_d, car en tant que groupe abélien, dI_t est exactement le noyau de l'application canonique $I_t \to \mathrm{Gal}(K_{nr}(x^{1/d})/K_{nr}) = \mu_d$. Ceci prouve que l'application qui à (a,d) associe θ_d^a est surjective, et par mise au même dénominateur, il apparaît que $\theta_d^a = \theta_{d'}^{a'}$ si et seulement si $a/d = a'/d' \mod \mathbb{Z}$. □

Si $\alpha \in \mathbb{Q}/\mathbb{Z}$ n'est pas d'ordre premier à p, il est possible de le projeter sur la partie de torsion première à p (c'est-à-dire $\mathbb{Z}_{(p)}/\mathbb{Z}$) via la somme directe

$$\mathbb{Q}/\mathbb{Z} = \mathbb{Z}[1/p]/\mathbb{Z} \oplus \mathbb{Z}_{(p)}/\mathbb{Z}.$$

On appelle alors caractère d'invariant α le caractère d'invariant α' où α' est la projection de α sur $\mathbb{Z}_{(p)}/\mathbb{Z}$, de sorte que l'application qui à α associe le caractère continu sur I_t d'invariant α est un morphisme de groupe surjectif de $\mathbb{Q}/\mathbb{Z}$ dans le groupe des caractères continu de I_t dans k_s^*.

En fait, on va se concentrer sur les « caractères fondamentaux », obtenus pour d de la forme $p^n - 1$.

Définition I.1.9 (Caractères fondamentaux)**.** Soit $n \geq 1$ un entier, $q = p^n$.
Un *caractère fondamental de niveau n* est un caractère obtenu en composant le caractère

$$\theta_{q-1} : I_t \to \mu_{q-1} = \mathbb{F}_q^* \subset k_s^*$$

avec un automorphisme de k_s^*. Autrement dit, les caractères fondamentaux de niveau n sont les n caractères

$$\theta_{q-1}, \theta_{q-1}^p, \cdots, \theta_{q-1}^{p^{n-1}}.$$

Ils sont d'invariants respectifs $\frac{p^i}{p^n-1}, i = 0, \cdots, n-1$.

Définition I.1.10. Pour tout $\alpha \in \mathbb{Q}$, on note

$$\mathfrak{m}_\alpha := \{x \in \overline{K}; v(x) \geq \alpha\}, \qquad \mathfrak{m}_\alpha^+ := \{x \in \overline{K}; v(x) > \alpha\},$$

et

$$V_\alpha := \mathfrak{m}_\alpha/\mathfrak{m}_\alpha^+.$$

La structure naturelle de $\overline{\mathcal{O}}$-module de $\mathfrak{m}_\alpha$ induit par passage au quotient une structure d'espace vectoriel de dimension 1 de V sur k_s, sur lequel G agit canoniquement.

La proposition suivante décrit entièrement l'action de $\mathrm{Gal}(\overline{K}/K)$ sur V_α.

Proposition I.1.11. *Soit $\sigma \in G$ et s l'image de σ dans $\mathrm{Gal}(k_s/k)$. Pour tout $\alpha \in \mathbb{Q}$, l'automorphisme de V_α défini par σ est s-linéaire. En particulier, pour tout $\sigma \in I$, l'automorphisme est k_s-linéaire, et l'action de I_t sur V_α est donnée par le caractère d'invariant α décrit par la proposition I.1.8.*

Démonstration. Pour tous $\lambda \in \overline{\mathcal{O}}$ et $v \in \mathfrak{m}_\alpha$, $\sigma(\lambda \cdot v) = \sigma(\lambda)\sigma(v)$ d'où la s-linéarité de l'automorphisme de V_α défini par σ. Cela induit automatiquement que celui-ci est k_s-linéaire pour $\sigma \in I$ par définition du groupe d'inertie, et comme I_p est un pro-p-groupe et que k_s^* ne contient aucun élément d'ordre p, I_p agit trivialement sur V_α d'où une action de I_t sur V_α. Reste à calculer son invariant en tant que caractère de I_t, qu'on note $\varphi_\alpha : I_t \to k_s^*$. L'application $\alpha \mapsto \varphi_\alpha$ est un morphisme de groupes car le produit sur $\overline{K}$ induit un isomorphisme de G-représentations $V_\alpha \otimes V_\beta \to V_{\alpha+\beta}$ pour tous $\alpha, \beta \in \mathbb{Q}$. Maintenant, pour $d \geq 1$ entier premier à p et x la racine d-ième d'une uniformisante de K, on a $\sigma(x) = \theta_d(x)x$ par définition de θ_d donc $\varphi_{1/d} = \theta_d$, donc $\varphi_{a/d}$ est d'invariant a/d pour tout $a \in \mathbb{Z}$.

Pour $\alpha \in \mathbb{Q}$ quelconque, il existe $q \in \mathbb{N}^*$ tel que $q\alpha = a/d$ avec d premier à p, et alors $\varphi_\alpha^q = \varphi_{a/d} = \theta_d^a$ qui est le caractère χ_α d'invariant α à la puissance q, donc $\varphi_\alpha = \chi_\alpha$ car k_s^* ne contient pas de q-torsion non triviale. □

Grâce à ces notations, on sait exactement comment agit I_t sur μ_p.

Corollaire I.1.3. *L'action naturelle de I_t sur μ_p est donnée par le caractère d'invariant $e/(p-1)$, avec e la ramification absolue de K.*

Démonstration. Soit $\alpha = e/(p-1) \in \mathbb{Q}$. Pour toute racine de l'unité $\zeta \in \mu_p$ différente de 1, $v(\zeta - 1) = \alpha$. En effet,

$$\prod_{\substack{\zeta \in \mu_p \\ \zeta \neq 1}} (\zeta - 1) = P(1)$$

avec $P = (1 - X^p)/(1 - X)$. Comme $P(1) = p$, sa valuation est e et chacun des $(\zeta - 1)$ a même valuation car ils sont tous conjugués, d'où $v(\zeta - 1) = \alpha$. L'application

$$\begin{array}{ccc} \mu_p & \longrightarrow & V_\alpha \\ \zeta & \longmapsto & \zeta - 1 \end{array}$$

est Galois-équivariante, et c'est un morphisme de groupes, car pour tous $\zeta, \zeta' \in \mu_p$,

$$\zeta\zeta' - 1 = \zeta(\zeta' - 1) + \zeta - 1 = \zeta' - 1 + \zeta - 1 \mod \mathfrak{m}_\alpha^+$$

car $\zeta = 1$ dans $\overline{k}$. Elle est injective car $v(\zeta - 1) = \alpha$ dès que $\zeta \neq 1$. On a donc une injection Galois-équivariante de groupes entre μ_p et V_α, et le résultat en découle car l'action de I_t sur V_α est donnée par le caractère d'invariant α d'après la proposition précédente. □

Soit maintenant E une courbe elliptique définie sur K et $\rho_{E,p} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(E[p])$ l'action naturelle de $\mathrm{Gal}(\overline{K}/K)$ sur la p-torsion. On suppose que la courbe elliptique E est semi-stable sur K, ce à quoi on peut se ramener en passant à une extension de K de degré 2,3,4 ou 6 grâce à l'algorithme de Tate (voir le détail de la preuve de la proposition VII.5.4 (*c*) de [Sil09]).

Les propositions suivantes résument les trois cas possibles.

Proposition I.1.12 (Cas ordinaire)**.**

Supposons que E a bonne réduction ordinaire.

(a) La droite de $E[p]$ se réduisant sur 0 modulo $\mathfrak{m}$ est stable par $\mathrm{Gal}(\overline{K}/K)$, donc $\rho_{E,p}$ est à valeurs dans un sous-groupe de Borel de $\mathrm{GL}(E[p])$, de type

$$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$

(b) L'image de I dans $\mathrm{GL}(E[p])$ est de la forme

$$\begin{pmatrix} \theta_{p-1}^e & * \\ 0 & 1 \end{pmatrix}.$$

(c) Si I_p agit trivialement sur $E[p]$, l'image de I dans $\mathrm{GL}(E[p])$ est un groupe cyclique d'ordre $(p-1)/\mathrm{pgcd}(p-1,e)$.

(d) Si I_p n'agit pas trivialement sur $E[p]$, l'image de I dans $\mathrm{GL}(E[p])$ est un groupe d'ordre $p(p-1)/\mathrm{pgcd}(p-1,e)$.

Démonstration. Le (a) et (b) sont directement donnés par la proposition 11 de [Ser72] et la discussion qui la précède. Pour le (c), si I_p agit trivialement sur $E[p]$, l'image de I ne contient pas de p-torsion, en conséquence une matrice de l'image de I est entièrement déterminée par ses coefficients diagonaux, c'est-à-dire par θ_{p-1}^e. On obtient donc un groupe cyclique d'ordre $(p-1)/\mathrm{pgcd}(p-1,e)$. Pour le (d), si I_p n'agit pas trivialement sur $E[p]$, il existe un élément d'ordre p dans l'image de I, qui ne peut être que de la forme $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ dans la base adaptée choisie, donc l'application qui à un élément de l'image de I associe son premier coefficient diagonal est surjective de noyau $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$, et d'après le (b), l'image de I est donc d'ordre $p(p-1)/\mathrm{pgcd}(p-1,e)$. □

Proposition I.1.13 (Cas multiplicatif)**.**

Supposons que E a réduction multiplicative.

(a) Il existe une $\mathbb{F}_p$-droite de $E[p]$ stable par I. De plus, l'image de I dans $\mathrm{GL}(E[p])$ est de la forme

$$\begin{pmatrix} \theta_{p-1}^e & * \\ 0 & 1 \end{pmatrix}.$$

(b) Si I_p agit trivialement sur $E[p]$, l'image de I dans $\mathrm{GL}(E[p])$ est un groupe cyclique d'ordre $(p-1)/\mathrm{pgcd}(p-1,e)$.

(c) Si I_p n'agit pas trivialement sur $E[p]$, l'image de I dans $\mathrm{GL}(E[p])$ est un groupe d'ordre $p(p-1)/\mathrm{pgcd}(p-1,e)$.

Démonstration. Le (a) vient de la proposition 13 de [Ser72], les (b) et (c) s'en déduisent comme dans la preuve précédente. □

Le cas le plus compliqué dans sa formulation est celui des courbes à réduction supersingulière. On en reproduit la preuve complète ici car la proposition n'est pas exactement écrite de cette manière dans [Ser72].

Proposition I.1.14 (Cas supersingulier)**.**

Supposons que E a bonne réduction supersingulière. Alors, $E[p]$ est isomorphe à un groupe formel à un paramètre sur K. Supposons aussi que $e < p+1$. Alors, deux cas sont possibles :

- *Le I-module $E[p]$ est simple (donc I_p agit trivialement) et il admet une structure de $\mathbb{F}_{p^2}$-espace vectoriel de rang 1 sur lequel I_t agit par le caractère $\theta_{p^2-1}^e$. En particulier, l'image de I_t dans $\mathrm{GL}(E[p])$ est un groupe cyclique de cardinal $(p^2-1)/\mathrm{pgcd}(p^2-1,e)$.*
- *Le I-module $E[p]$ a un sous-module de dimension 1, et l'action de I sur $E[p]$ est alors de la forme*

$$\begin{pmatrix} \theta_{p-1}^a & * \\ 0 & \theta_{p-1}^b \end{pmatrix}$$

avec $0 \leq a, b \leq e$ tels que $a + b = e$. De plus, I_p n'agit pas trivialement sur $E[p]$, donc le cardinal de l'image de I dans $\mathrm{GL}(E[p])$ est divisible par p.

Démonstration. Comme E a bonne réduction supersingulière, le module galoisien $E[p]$ est isomorphe à l'ensemble des points de p-torsion d'une certaine loi de groupe formel $F(X, Y)$ sur $\mathcal{O}_K[[X, Y]]$, dont la réduction sur $k[[X, Y]]$ est de hauteur 2 ([Sil09], Théorème V.3.1). Il nous suffit donc d'étudier ces points. Considérons le polygone de Newton de la série formelle

$$[p](X) = \sum_{i=1}^{+\infty} a_i X^i \qquad (a_i \in \mathcal{O}_K)$$

de multiplication par p pour la loi du groupe formel. On sait que $a_1 = p$ et comme le groupe formel est de hauteur 2, le premier coefficient inversible de la série est a_{p^2}. Notons V l'ensemble des zéros de $[p]$ dans $\mathfrak{m}$, c'est un espace vectoriel de rang 2 sur $\mathbb{F}_p$. Pour l'étudier plus précisément, considérerons le polygone de Newton de $[p]$. Deux cas se présentent à nous.

• Le polygone de Newton entre 1 et p^2 est une ligne droite du point $(1, e = v(p))$ au point $(p^2, 0 = v(a_{p^2}))$:

Les zéros de la série $[p]$ sont alors tous de valuation $\alpha = \frac{e}{p^2-1}$ qui est l'opposé de la pente de cette ligne droite. En conséquence, le groupe V s'injecte dans V_α par un homomorphisme Galois-équivariant. Le groupe I_p agit donc trivialement sur V car il opère trivialement sur V_α, et I_t agit donc sur V via le caractère $\theta_{p^2-1}^e$. Or, l'image de θ_{p^2-1} dans k_s^* est $\mathbb{F}_{p^2}^*$. On peut donc munir V d'une structure de $\mathbb{F}_{p^2}$-espace vectoriel de rang 1, pour laquelle I_t agit par automorphismes $\mathbb{F}_{p^2}$-linéaires. Une telle copie de $\mathbb{F}_{p^2}^*$ dans $\mathrm{GL}(V)$ est constitués de $p-1$ homothéties et de p^2-p éléments n'ayant aucune droite fixe. Or, pour $e < p+1$, $p^2-1/\mathrm{pgcd}(p^2-1, e) > p-1$, donc il existe au moins un automorphisme de V provenant de G n'ayant pas de droite propre : en conséquence, le I-module $E[p]$ est simple, et I_t agit sur $E[p]$ via le caractère $\theta_{p^2-1}^e$.

• Le polygone de Newton entre 1 et p^2 est une ligne brisée, avec un point intermédiaire (x, a) : Tout d'abord, l'ensemble des points nuls ou de valuation au moins $\alpha = \frac{e-a}{x-1}$ forme un sous-groupe non trivial V_0 de V de cardinal x, or V est de cardinal p^2, donc $x = p$. On a donc une sous-$\mathbb{F}_p$-droite de V constituée de 0 et d'éléments de valuation $\alpha = \frac{e-a}{p-1}$. Le reste de V est, d'après le polygone de Newton, constitué d'éléments de valuation $\beta = \frac{a}{p^2-p}$. Or, comme $a < e < p+1$, chacun de ces éléments engendre une extension sauvagement ramifiée de K puisque leur valuation a un dénominateur divisible par p. Il est donc impossible que I_p agisse trivialement sur ces éléments (par définition, si I_p agissait trivialement, ces éléments seraient dans K_t). En conséquence, I_p n'agit pas trivialement sur $E[p]$, donc l'image de I dans $\mathrm{GL}(E[p])$ a une p-partie. De plus, I agit sur V_α via le caractère θ_{p-1}^a et le déterminant de l'action de I sur $E[p]$ est exactement le caractère cyclotomique donné par l'action de I sur μ_p, c'est-à-dire θ_{p-1}^e, d'où la forme de l'action de I sur $E[p]$. □

Ceci permet de conclure sur le cas exceptionnel avec la proposition suivante, dont les $\mathbb{Q}$-courbes ne seront qu'un cas particulier.

Proposition I.1.15. *Soit K un corps de nombres et E une courbe elliptique sur K. Pour tout $p > 30[K : \mathbb{Q}] + 1$, l'image de $\mathbb{P}\rho_{E,p} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{PGL}(E[p])$ n'est pas contenue dans un sous-groupe exceptionnel (de la forme $\mathcal{A}_4$, $\mathcal{A}_5$ ou $\mathcal{S}_4$). En particulier, pour une $\mathbb{Q}$-courbe E définie sur le corps quadratique K, l'image de $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas contenue dans un sous-groupe exceptionnel pour $p > 67$.*

Démonstration. Soit E une courbe elliptique sur le corps de nombres K. Soit $\mathfrak{p}$ un idéal premier de $\mathcal{O}_K$ au-dessus de p de degré de ramification e. Considérons le complété $K_\mathfrak{p}$ de K pour la valuation associée à $\mathfrak{p}$. Il existe par l'algorithme de Tate une extension K' de $K_\mathfrak{p}$ de degré de ramification $d = 1, 2, 3, 4$ ou 6 telle que $E_{K'}$ est semi-stable, notons I' le groupe d'inertie associé à K'. Si I'_p agit non trivialement sur $E[p]$, l'image de $\mathrm{PGL}(E[p])$ contient un p-groupe et n'est donc pas exceptionnelle dès que $p > 5$. Sinon, les bornes données dans chacun des trois cas permettent de dire que l'image de I'_t par $\mathbb{P}\overline{\rho}_{E,p}$ contient un sous-groupe cyclique d'ordre au moins

$(p-1)/\mathrm{pgcd}(p-1, ed)$, qui est strictement supérieur à 5 lorsque $p > 5ed+1$, et dans ce cas cette image ne peut être incluse dans un sous-groupe exceptionnel. Comme $d \leq 6$ et $e \leq [K:\mathbb{Q}]$, on en déduit la borne de la proposition. □

Le reste de la section sera consacré à formaliser les trois cas restants (Borel, normalisateur de Cartan déployé, normalisateur de Cartan non déployé) en termes de courbes modulaires.

I.1.4 Traduction du problème sur des courbes modulaires

On suppose désormais que l'image de $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas contenu dans un sous-groupe exceptionnel ni dans $\mathrm{PSL}(E[p])$, et on fixe K un corps quadratique dans tout ce paragraphe.

Pour tout nombre premier p et tout entier $N \geq 1$:

• La courbe modulaire $X_0(N)_{\mathbb{Z}}$ est le schéma de modules grossier compactifié sur $\mathbb{Z}$ paramétrant les classes d'isomorphisme de couples (E, C_N) avec E une courbe elliptique et C_N un sous-groupe de N-torsion cyclique de E. Sa fibre générique est la courbe modulaire sur $\mathbb{Q}$ correspondant au sous-groupe de congruences

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),\, c \equiv 0 \mod N \right\}.$$

• La courbe modulaire $X_{\mathrm{split}}(p)_{\mathbb{Z}}$ est le schéma de modules grossier compactifié sur $\mathbb{Z}$ paramétrant les classes d'isomorphisme de couples $(E, \{A_p, B_p\})$ avec E une courbe elliptique et $\{A_p, B_p\}$ une paire non distinguée de deux sous-groupes cycliques d'ordre p distincts de E. Sa fibre générique est la courbe modulaire sur $\mathbb{Q}$ correspondant au sous-groupe de congruences

$$\Gamma_{\mathrm{split}}(p) := \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}), \gamma \equiv \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \text{ ou } \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix} \mod \mathrm{p} \right\}.$$

• La courbe modulaire $X_{\mathrm{nonsplit}}(p)$ est le schéma de modules grossier compactifié sur $\mathbb{Z}$ paramétrant les classes d'isomorphisme de couples $(E, \{D, D'\})$ avec E une courbe elliptique et $\{D, D'\}$ une paire de $\mathbb{F}_{p^2}$-droites conjuguées de $E[p] \otimes_{\mathbb{F}_p} \mathbb{F}_{p^2}$. Sa fibre générique sur $\mathbb{Q}$ est la courbe modulaire sur $\mathbb{Q}$ correspondant à un sous-groupe de congruences

$$\Gamma_{\mathrm{nonsplit}}(p) := \{\gamma \in \mathrm{SL}_2(\mathbb{Z}), \gamma \equiv \alpha \mod p, \alpha \in G\},$$

où G est le normalisateur d'un sous-groupe de Cartan non déployé fixé de $\mathrm{GL}_2(\mathbb{F}_p)$. Les définitions et propositions suivantes sont tirées de [Ell04].

Définition I.1.16. Soit d un entier sans facteur carré premier à p. On définit les courbes modulaires

$$X_0^{\mathrm{s}}(d;p) := X_0(d) \times_{X(1)} X_{\mathrm{split}}(p)$$

$$X_0^{\mathrm{ns}}(d;p) := X_0(d) \times_{X(1)} X_{\mathrm{nonsplit}}(p).$$

Alors, par construction :

• La courbe modulaire $X_0(dp)_{\mathbb{Z}}$ (qui est le produit fibré $X_0(d) \times_{X(1)} X_0(p)$ muni des morphismes d'oubli) est l'espace de modules grossier compactifié sur $\mathbb{Z}$ des classes d'isomorphisme de triplets (E, C_d, C_p) avec E une courbe elliptique, C_d un sous-groupe cyclique d'ordre d de E et C_p un sous-groupe cyclique d'ordre p de E.

• La courbe modulaire $X_0^{\mathrm{s}}(d;p)_{\mathbb{Z}}$ est l'espace de modules grossier compactifié sur $\mathbb{Z}$ des classes d'isomorphisme de triplets $(E, C_d, \{A_p, B_p\})$ avec E une courbe elliptique, C_d un sous-groupe cyclique d'ordre d de E et A_p, B_p deux sous-groupes d'ordre p distincts de E.

• La courbe modulaire $X_0^{\mathrm{ns}}(d;p)_{\mathbb{Z}}$ est l'espace de modules grossier compactifié sur $\mathbb{Z}$ des classes d'isomorphisme de triplets $(E, C_d, \{D, D'\})$ avec E une courbe elliptique, C_d un sous-groupe cyclique d'ordre d de E et $\{D, D'\}$ une paire de $\mathbb{F}_{p^2}$-droites conjuguées de $E[p] \otimes_{\mathbb{F}_p} \mathbb{F}_{p^2}$.

Les courbes modulaires $X_0(dp)$, $X_0^{\mathrm{s}}(d;p)$ et $X_0^{\mathrm{ns}}(d;p)$ sont toutes les trois munies d'une involution w_d qui s'exprime sur les classes de représentants comme suit :

$$\begin{array}{rcccc} w_d(E,C_d,C_p) & = & (E/C_d, E[d]/C_d, C_p/C_d) & \text{sur} & X_0(dp). \\ w_d(E,C_d,\{A_p,B_p\}) & = & (E/C_d, E[d]/C_d, \{A_p/C_d, B_p/C_d\}) & \text{sur} & X_0^{\mathrm{s}}(d;p). \\ w_d(E,C_d,\{D,D'\}) & = & (E/C_d, E[d]/C_d, \{D/C_d, D'/C_d\}) & \text{sur} & X_0^{\mathrm{ns}}(d;p) \end{array}$$

où la notation H/C_d pour H un sous-groupe de E est l'image de H par une isogénie $E \to E/C_d$ de noyau C_d.

Pour chacune de ces courbes X, on note X^K la courbe tordue de X par w_d et $\mathrm{Gal}(K/\mathbb{Q})$, c'est-à-dire la courbe admettant un isomorphisme $\phi : X \to X^K$ défini sur K tel que pour σ l'automorphisme de K, ${}^\sigma\phi = w_d \circ \phi$. Ses points rationnels sont alors mis en bijection par ϕ avec les points $P \in X(K)$ tels que ${}^\sigma P = w_d P$. Les $\mathbb{Q}$-courbes nous fournissent exactement de tels points.

Proposition I.1.17. *Soit K un corps quadratique. Soit d un entier sans facteur carré, E une $\mathbb{Q}$-courbe de degré d sans multiplication complexe définie sur K, et p un nombre premier ne divisant pas d.*

Si l'image de $\mathbb{P}\overline{\rho}_{E,p}$ est incluse dans un sous-groupe de Borel de $\mathrm{PGL}(E[p])$ *(resp. le normalisateur d'un sous-groupe de Cartan déployé, le normalisateur d'un sous-groupe de Cartan non déployé), alors on peut canoniquement associer à E un point de $X_0(dp)^K(\mathbb{Q})$ (resp. $X_0^{\mathrm{s}}(m;p)(\mathbb{Q})$, $X_0^{\mathrm{s}}(m;p)(\mathbb{Q})$).*

Démonstration. Notons $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ l'automorphisme non trivial du corps K. Par hypothèse, on a une isogénie $\mu : E \to {}^\sigma E$ de degré d de noyau cyclique, noté C_d et $\widehat{\mu}$ l'isogénie duale. Soit maintenant $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Si $\tau_{|K} = \mathrm{Id}_K$, ${}^\tau\mu$ est une isogénie de degré d de E vers ${}^\sigma E$, ce ne peut donc être que $\pm\widehat{\mu}$ car ${}^\sigma E$ et E sont sans multiplication complexe. En particulier, ${}^\tau C_d = C_d$, ce sous-groupe cyclique est donc défini sur K. Ensuite, si $\tau_{|K} = \sigma$, ${}^\tau\mu$ est une isogénie de ${}^\sigma E$ vers E de degré d, c'est donc $\pm\hat{\mu}$ pour les mêmes raisons. En particulier, ${}^\tau C_d = \ker\hat{\mu} = E[d]/C_d$ (on note par la suite ${}^\sigma C_d$ ce sous-groupe de ${}^\sigma E$).

Nous allons traiter le cas des sous-groupes de Borel, les deux autres sont similaires. Soit C_p le sous-groupe cyclique d'ordre p de E fixé par la représentation projective. Alors, pour tout $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, si $\tau_{|K} = \mathrm{Id}_K$, ${}^\tau C_p = C_p$ et si $\tau_{|K} = \sigma$, $\widehat{\mu}({}^\tau C_p) = C_p$. Ainsi, C_p est stable par $\mathrm{Gal}(\overline{K}/K)$ et ${}^\sigma C_p$ est bien défini et vérifie $\widehat{\mu}({}^\sigma C_p) = C_p$ par hypothèse, d'où $\sigma C_p = \mu(C_p)$.

Alors,

$${}^\sigma(E,C_d,C_p) = ({}^\sigma E, {}^\sigma C_d, {}^\sigma C_p) = (E/C_d, E[d]/C_d, C_p + C_d/C_d) = w_d(E,C_d,C_p).$$

Le triplet (E,C_d,C_p) est donc un point $\mathbb{Q}$-rationnel de la courbe tordue $X_0(dp)^K$. □

La démarche de notre preuve est la suivante : montrer que pour un nombre premier assez grand p et un entier $d > 1$, aucune des courbes modulaires tordues $X_0(dp)^K, X_0^{\mathrm{s}}(d;p)^K$ et $X_0^{\mathrm{ns}}(d;p)^K$ n'admet de point rationnel qui ne soit ni une pointe ni un point CM.

Pour cela, le début du travail consiste à montrer qu'un tel point aurait bonne réduction modulo tout idéal premier de $\mathcal{O}_K$, selon l'approche initiée par Mazur [Maz77], décrite dans la section suivante (et appliquée dans les trois cas dans les sections I.4, I.5 et I.6).

I.2 La méthode de Mazur

I.2.1 Rappels schématiques

Modèles et réduction

Commençons par quelques rappels sur les schémas.

Définition I.2.1. Soit $f : X \to Y$ un morphisme de schémas.

Tout point y de Y correspond à un morphisme de schémas de $\mathrm{Spec}(k(y)) \longrightarrow Y$ avec $k(y)$ le corps résiduel de Y en y. La *fibre de f en y* est le produit fibré

$$X_y := X \times_Y \mathrm{Spec}(k(y)).$$

L'espace topologique sous-jacent est canoniquement homéomorphe à $f^{-1}(y)$. En conséquence, pour tout point $x \in f^{-1}(y)$, on notera x_y le point de X_y correspondant. Dans le cas où Y est irréductible de point générique η, on appelle X_η la *fibre générique* du morphisme. Si y est un point fermé de Y, on appelle X_y la *fibre spéciale* en y. Une *famille de schémas* sur Y est l'ensemble des fibres d'un morphisme de schémas $f : X \to Y$.

Remarque I.2.1. Les notations ci-dessus font plutôt référence à X qu'à f, alors que leur définition rigoureuse implique l'usage du morphisme : dans la plupart des cas, le morphisme structural sera évident, de sorte qu'on préfère reporter les notations sur X.

Définition I.2.2 (Modèle)**.**
Soit K un corps et $\mathcal{O}_K$ un anneau intègre de corps des fractions K. Si X est un schéma sur K, un *modèle* de X est un schéma $\mathcal{X}$ sur $\mathcal{O}_K$ dont la fibre générique est isomorphe au schéma X.

Il existe de nombreux modèles différents d'un même schéma en général, mais on cherche à ce qu'ils aient de bonnes propriétés.

Définition I.2.3 (Morphisme propre)**.**
Un morphisme de schémas $f : X \to Y$ est *propre* s'il est de type fini, séparé et universellement fermé, c'est-à-dire que pour tout Y-schéma Y', le morphisme de schémas $X' = X \times_Y Y' \to Y'$ obtenu par changement de base est topologiquement fermé.

Les modèles propres forment une classe intéressantes de modèles pour de multiples raisons ([Mum99], § II.7 ou [Liu02], section 3.3.2).
On note, lorsque X et Y sont deux Z-schémas, $X_Z(Y)$ l'ensemble des morphismes de Z-schémas de X vers Y. La proposition suivante est l'extension aux schémas de base de Dedekind d'une propriété classique des schémas propres ([Liu02], Corollaire 3.3.26).

Proposition I.2.4 (Extension des points sur la fibre générique)**.**
Soit $f : X \to Y$ un morphisme propre. Pour tout Y-schéma de Dedekind S de corps des fractions K, l'application canonique $X_Y(S) \to X_Y(K)$ est bijective. En d'autres mots, à tout Y-point K-rationnel x de X, on peut associer de manière unique un Y-morphisme $x_S : S \to X$ dont la restriction à la fibre générique de S est x. De plus, pour tout $s \in S$, la section $\mathrm{Spec}\,\mathcal{O}_{S,s} \to S \to X$ *est exactement le morphisme obtenu par le critère valuatif appliqué à $\mathcal{O}_{S,s}$ et x.*

Un autre type de modèle intéressant est le modèle de Néron.

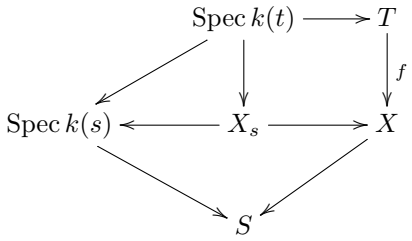
Définition I.2.5 (Modèle de Néron)**.**
Soit S un schéma de Dedekind de corps des fractions K. Soit V une variété définie sur K. Un schéma $\mathcal{V} \to S$ est un modèle de Néron de V/K s'il est lisse sur S et que pour tout schéma lisse $\mathcal{X} \to S$ de fibre générique X/K, tout morphisme de K-schémas $f : X \to V$ s'étend de manière unique en un morphisme de S-schémas $\mathcal{X} \to \mathcal{V}$.

Le lemme de Yoneda prouve comme d'habitude qu'un modèle de Néron, s'il existe, est unique à unique isomorphisme près. L'existence d'un tel modèle est un problème difficile, mais Néron lui-même a prouvé le résultat suivant [Nér64].

Proposition I.2.6. *Soit A une variété abélienne sur K. Alors il existe un modèle de Néron $\mathcal{A} \to S$ de A sur S, qui est également un schéma en groupes sur S.*

Passons maintenant à la réduction.

Définition I.2.7. Soient X et T deux S-schémas où S et T sont deux schémas de Dedekind, et $f : T \to X$ un S-morphisme de schémas. Soient $s \in S$ et $t \in T$ au-dessus de s. La *réduction de f modulo t* est le point $f(t)$ vu dans X_s. Plus précisément, c'est le morphisme de S-schémas $\operatorname{Spec} k(t) \to X_s$ défini par le diagramme commutatif suivant, où toutes les autres flèches sont canoniques :



C'est donc un $k(t)$-point de X_s.

Si $\mathcal{X}$ est un S-schéma propre de fibre générique X et T un S-schéma de Dedekind de corps des fractions K, pour tout point $x \in X(K)$, on appelle *réduction modulo $t \in T$ de x* la réduction modulo t de l'unique S-morphisme $T \to \mathcal{X}$ valant x sur la fibre générique, qu'on identifiera souvent à un point de X_s où s est le point de S au-dessous de t.

Remarque I.2.2. Si T' est un schéma de Dedekind au-dessus de T de corps des fractions L (qui est donc une extension de K), tout point K-rationnel de $\mathcal{X}$ est en particulier L-rationnel, on a donc un morphisme $T' \to \mathcal{X}$, et celui-ci se factorise par $T' \to T \to \mathcal{X}$. Ainsi, la réduction modulo $t' \in T'$ du point K-rationnel est la même que sa réduction modulo $t \in T$ où t est le point de T au-dessous de t' dans T.

Espaces cotangents et immersions formelles

Définition I.2.8 (Immersion formelle)**.**

Soit $f : X \to Y$ un morphisme de schémas, x un point de X d'image $y = f(x)$. Le morphisme f est une *immersion formelle en x* si l'homomorphisme $\widehat{f_x} : \widehat{\mathcal{O}}_{Y,y} \to \widehat{\mathcal{O}}_{X,x}$ déduit de f par passage aux complétés des anneaux locaux est surjectif.

Avant de comprendre en quoi une telle définition est intéressante, nous allons commencer par en rappeler une caractérisation classique, fondamentale pour la suite.

Définition I.2.9 (Espace cotangent)**.**

Soit X un schéma, et $x \in X$. L'*espace cotangent de X en x* est le $k(x)$-espace vectoriel $\mathfrak{m}_{X,x}/\mathfrak{m}_{X,x}^2$.

Proposition I.2.10 (Caractérisation des immersions formelles)**.**

Soient X, Y deux schémas localement noethériens, $f : X \to Y$ un morphisme de schémas, x un point de X et $y = f(x)$. On note $\mathfrak{m}_x$ et $\mathfrak{m}_y$ les idéaux maximaux respectifs de $\mathcal{O}_{X,x}$ et $\mathcal{O}_{Y,y}$. Alors, f est une immersion formelle en x si et seulement si les deux conditions suivantes sont vérifiées :

- *L'application $k(y) \to k(x)$ induite par f sur les corps résiduels est un isomorphisme.*
- *L'application cotangente $\mathfrak{m}_y/\mathfrak{m}_y^2 \to \mathfrak{m}_x/\mathfrak{m}_x^2$ induite par f sur les espaces cotangents est surjective.*

Remarque I.2.3. Vu son caractère local, cette caractérisation est en fait un résultat d'algèbre commutative : étant donné un homomorphisme d'anneaux noethériens locaux $\varphi : A \to B$, montrer que l'homomorphisme $\widehat{\varphi} : \widehat{A} \to \widehat{B}$ induit sur les anneaux complétés est surjectif si et seulement si les homomorphismes induits $A/\mathfrak{m}_A \to B/\mathfrak{m}_B$ sur les corps résiduels et $\mathfrak{m}_A/\mathfrak{m}_A^2 \to \mathfrak{m}_B/\mathfrak{m}_B^2$ sur les espaces cotangents sont surjectifs. La preuve de ce résultat est exposée juste après cette remarque, les lecteurs peu désireux de la lire peuvent d'emblée passer à l'interprétation géométrique des immersions formelles (proposition I.2.13).

Démontrons tout d'abord le résultat suivant, qui est un avatar du lemme de Nakayama pour les anneaux complets, qui ne suppose pas que le module est de type fini, mais permet au contraire de le démontrer.

Lemme I.2.11. *Soient A un anneau local complet d'idéal maximal $\mathfrak{m}$ et M un A-module tel que $\cap_k \mathfrak{m}^k M = (0)$. Si les classes de $m_1, \dots, m_r \in M$ modulo $\mathfrak{m}M$ engendrent $M/\mathfrak{m}M$ sur $A/\mathfrak{m}$, alors $m_1, \dots, m_r$ engendrent M sur A.*

Démonstration. Soient $m_1, \dots, m_r \in M$ tels que $\overline{m_1}, \dots, \overline{m_r}$ engendrent $M/\mathfrak{m}M$. On a donc

$$M = \mathfrak{m}M + \sum_{i=1}^{r} A.m_i \tag{I.2}$$

Démontrons par récurrence que pour tout $k \in \mathbb{N}$, on peut écrire

$$m = \sum_{i=1}^{r} \left(\sum_{j=0}^{k} a_i^{(j)} \right) m_i + m'_{k+1}$$

avec $a_i^{(j)} \in \mathfrak{m}^j$ pour tous i, j et $m'_{k+1} \in \mathfrak{m}^{k+1}M$.

Pour $k = 0$, c'est exactement la formule (I.2). Ensuite, si l'assertion est vraie pour $k \in \mathbb{N}^*$, si m est écrit sous la forme précédente, on peut écrire

$$m'_{k+1} = \sum_{\ell} s_\ell n_\ell$$

avec $s_\ell \in \mathfrak{m}^{k+1}$ et $n_\ell \in M$ pour tout ℓ. Alors, on écrit grâce à la formule (I.2) chaque n_ℓ sous la forme

$$n_\ell = \sum_{i=1}^{r} t_{\ell,i} m_i + n'_\ell$$

avec $t_{\ell,i} \in A$ et $n'_\ell \in \mathfrak{m}.M$, et on regroupe ceci dans l'écriture de m pour obtenir

$$m = \sum_{i=1}^{r} \Big(\sum_{j=0}^{k} a_i^{(j)} + \sum_{\ell} t_{\ell,i} s_\ell \Big) m_i + \sum_{\ell} s_\ell n'_\ell.$$

et on pose alors $a_i^{(k+1)} = \sum_\ell t_{\ell,i} s_\ell \in \mathfrak{m}^{k+1}$ et $m'_{k+2} = \sum_\ell s_\ell n'_\ell \in \mathfrak{m}^{k+2}.M$, ce qui prouve la récurrence.

Pour tout $1 \leq i \leq r$, la série $\sum_{j=0}^{k} a_i^{(j)}$ converge dans A car celui-ci est complet et la valuation $\mathfrak{m}$-adique du terme général tend vers l'infini. Soit $a_i \in A$ une limite de la série (on peut en avoir plusieurs si A n'est pas séparé). Pour tout $j \in \mathbb{N}$, $a_i^{(j)} \in \mathfrak{m}^j$, on a $a_i - \sum_{j=0}^{k} a_i^{(j)} \in \mathfrak{m}^{k+1}$ pour tout $k \in \mathbb{N}$, ce qui nous permet d'écrire que

$$m - \sum_{i=1}^{r} a_i m_i = m'_{k+1} - \sum_{i=1}^{r} (a_i - \sum_{j=0}^{k} a_i^{(j)}) m_i \in \mathfrak{m}^{k+1} M$$

et ce pour tout $k \in \mathbb{N}$, donc $m = \sum_{i=1}^{r} a_i m_i$ par hypothèse sur M, donc le A-module M est bien engendré par $m_1, \dots, m_r$. □

Passons maintenant à la preuve de la caractérisation des immersions formelles.

Démonstration. Rappelons que $A/\mathfrak{m}_A \cong \widehat{A}/\mathfrak{m}_{\widehat{A}}$ et de même pour B, de sorte que le morphisme induit par f sur les corps résiduels est le même que celui induit par $\hat{f}$ sur les corps résiduels de $\widehat{A}$ et $\widehat{B}$ via ces isomorphismes naturels. Ainsi, il est surjectif si et seulement si

$$\widehat{B} = \mathfrak{m}_{\widehat{B}} + \hat{f}(\widehat{A}) = \mathfrak{m}_{\widehat{B}} + \widehat{A}.1 \tag{I.3}$$

Supposons d'abord que $\hat{f}$ est surjective. L'égalité $(I.3)$ est évidente, et donc le morphisme de corps résiduels est un isomorphisme. Ensuite, pour tout $\overline{b} \in \mathfrak{m}_B/\mathfrak{m}_B^2$, soit $b \in \mathfrak{m}_B$ un relèvement de $\overline{b}$. L'image naturelle de b dans $\widehat{B}$ est l'image par $\hat{f}$ d'une suite cohérente $(\overline{a_n}) \in \widehat{A}$. En particulier, par définition de $\hat{f}$, on a $b = f(a_2) \mod \mathfrak{m}_B^2$, et comme $b \in \mathfrak{m}_B$, on a forcément $a_2 \in \mathfrak{m}_A$ car il n'est pas inversible dans B. Ainsi, $\overline{b}$ a bien un antécédent dans $\mathfrak{m}_A/\mathfrak{m}_A^2$, c'est-à-dire que l'application induite est surjective.

Réciproquement, supposons que le morphisme de corps résiduels est un isomorphisme et que le morphisme d'espaces cotangents est surjectif. On a $(I.3)$ et $\mathfrak{m}_B = \mathfrak{m}_B^2 + f(\mathfrak{m}_A)$, d'où en tant que $\widehat{B}$-modules, on a

$$\mathfrak{m}_B\widehat{B} = \mathfrak{m}_B^2\widehat{B} + f(\mathfrak{m}_A)\widehat{B}.$$

Or, on a $\mathfrak{m}_B\widehat{B} = \mathfrak{m}_{\widehat{B}}$ et $\mathfrak{m}_B^2\widehat{B} = \mathfrak{m}_{\widehat{B}}^2$ car B est noethérien ([AM94], Proposition 10.15), donc l'égalité se réécrit

$$\mathfrak{m}_{\widehat{B}} = \mathfrak{m}_{\widehat{B}}^2 + f(\mathfrak{m}_A)\widehat{B}$$

et comme $\widehat{B}$ est noethérien ([AM94], Théorème 10.26), $\mathfrak{m}_{\widehat{B}} = f(\mathfrak{m}_A)\widehat{B}$ par le lemme de Nakayama classique, c'est-à-dire que $\mathfrak{m}_{\widehat{B}} = \mathfrak{m}_A.\widehat{B}$ en tant que A-module, donc $\mathfrak{m}_{\widehat{B}} = \mathfrak{m}_{\widehat{A}}\widehat{B}$ car $\mathfrak{m}_{\widehat{A}} = \mathfrak{m}_A\widehat{A}$, A étant également noethérien.

Enfin, d'après le lemme I.2.11 appliqué à l'anneau $\widehat{A}$ et au $\widehat{A}$-module $\widehat{B}$ (car $\cap_k \mathfrak{m}_{\widehat{B}}^k = 0$ par le lemme de Nakayama appliqué à cette intersection d'idéaux et à $\widehat{B}$), comme $\widehat{B}/\mathfrak{m}_{\widehat{A}}.\widehat{B} = \widehat{B}/\mathfrak{m}_{\widehat{B}}$ est un $B/\mathfrak{m}_B = A/\mathfrak{m}_A$-espace vectoriel engendré par 1, $\widehat{B}$ est un $\widehat{A}$-module engendré par 1, c'est-à-dire que $\hat{f}$ est surjective. □

Corollaire I.2.1. *Soit $f : X \to Y$ un morphisme de S-schémas localement noethériens. Alors, f est une immersion formelle en un point $x \in X$ au-dessus de $s \in S$ si et seulement si l'application fibrée $f_s : X_s \to Y_s$ est une immersion formelle en x_s.*

Démonstration. Par localité du problème, on peut prendre des voisinages affines de x, y et s et les anneaux locaux associés, de sorte qu'il suffit de prouver un résultat d'algèbre commutative. Plus précisément, soit A, B, C sont des anneaux locaux noethériens d'idéaux maximaux respectifs $\mathfrak{m}_A$, $\mathfrak{m}_B$ et $\mathfrak{m}_C$ avec B et C des A-algèbres locales et $f : B \to C$ un morphisme de A-algèbres locales. Il suffit de montrer que $\widehat{f}$ est surjective si et seulement si $\widehat{f_A} : \widehat{B_A} \to \widehat{C_A}$ l'est, avec $B_A := B \otimes_A A/\mathfrak{m}_A$, $C_A := C \otimes_A A/\mathfrak{m}_A$ et $f_A : B_A \to C_A$ le morphisme induit par f. Dans le sens direct, les surjections évidentes $\widehat{B} \to \widehat{B_A}$ et $\widehat{C} \to \widehat{C_A}$ s'insèrent dans le diagramme commutatif

$$\begin{array}{ccc} \widehat{B} & \xrightarrow{\widehat{f}} & \widehat{C} \\ \downarrow & & \downarrow \\ \widehat{B_A} & \xrightarrow{\widehat{f_A}} & \widehat{C_A} \end{array}$$

donc si $\widehat{f}$ est surjective, $\widehat{f_A}$ l'est. Réciproquement, si $\widehat{f_A}$ est surjective, le morphisme de corps résiduels est un isomorphisme ce qui prouve que $B/\mathfrak{m}_B \to C/\mathfrak{m}_C$ l'est (on voit directement que les corps résiduels de B_A et B, resp. C_A et C sont les mêmes). Ensuite, le morphisme d'espaces cotangents est surjectif, c'est-à-dire que

$$(\mathfrak{m}_C/\mathfrak{m}_A C) = (\mathfrak{m}_C/\mathfrak{m}_A C)^2 + f(\mathfrak{m}_B)/f(\mathfrak{m}_A B)$$

donc

$$\mathfrak{m}_C = \mathfrak{m}_C^2 + f(\mathfrak{m}_B) + \mathfrak{m}_A C = \mathfrak{m}_C^2 + f(\mathfrak{m}_B)C.$$

Á partir de là, on peut multiplier par $\widehat{C}$ et reprendre la fin de la preuve précédente pour établir que $\widehat{f}$ est surjective.

□

Définition I.2.12 (Espace cotangent le long d'une section)**.**
Soit X un S-schéma séparé et $s : S \to X$ une section de X. *L'espace cotangent de X le long de la section s* est le $\mathcal{O}_S$-module

$$\mathrm{Cot}_s(X) = s^*(\Omega_{X/S}),$$

où $\Omega_{X/S}$ est le faisceau des différentielles relatives de X sur S ([Liu02], Proposition 6.1.17). Dans le cas où $S = \mathrm{Spec}\, k$, une section $s : S \to X$ est un point x de X de corps résiduel $k(x) = k$. Alors, $s^*(\Omega_{X/k}) = \Omega_{\mathcal{O}_{X,x}/k} \otimes_k k(x) = \mathfrak{m}_x/\mathfrak{m}_x^2$, ce qui correspond donc bien à la définition usuelle.

Pour plus de résultats sur les espaces cotangents et les faisceaux de différentielles relatives, voir le chapitre 6 de [Liu02].

L'interprétation géométrique de l'immersion formelle est la suivante. Étant donné un point x d'un schéma X, on appelle *voisinage formel de x* le morphisme $\mathrm{Spec}\,\widehat{\mathcal{O}_{X,x}} \to X$ canoniquement associé à x. Considérons maintenant une immersion formelle $f : X \to Y$ en x, et $y = f(x)$. L'homomorphisme d'anneaux locaux complétés $\widehat{f}_x : \widehat{\mathcal{O}_{Y,y}} \to \widehat{\mathcal{O}_{X,x}}$ se traduit canoniquement en un morphisme entre les voisinages formels $\mathrm{Spec}\,\widehat{\mathcal{O}_{X,x}} \to \mathrm{Spec}\,\widehat{\mathcal{O}_{Y,y}}$, qu'on considère comme la « restriction » de f aux voisinages formels de x et y. Vue sur ces voisinages, la surjectivité de $\widehat{f}_x$ s'interprète alors comme le fait que cette restriction est une immersion fermée. Dire que f est une immersion formelle en x est donc dire que c'est une immersion sur le « voisinage formel » de x. Or, en géométrie différentielle, une propriété naturelle d'une immersion en un point est qu'elle empêche deux sections distinctes transverses en x d'être égales par composition. Cette propriété se transpose au langage des schémas comme suit.

Proposition I.2.13 (Propriété des immersions formelles)**.**
Soit X un schéma séparé et $f : X \to Y$ une immersion formelle en $x \in X$. Soient T un schéma intègre noethérien, t un point de T et g, h deux morphismes de T dans X tels que $g(t) = h(t) = x$ et $f \circ g = f \circ h$. Alors, $g = h$.

Remarque I.2.4. Les grandes étapes de la démonstration peuvent se concevoir de façon géométrique : tout d'abord, on montre que g et h sont égaux du voisinage infinitésimal de t vers le voisinage infinitésimal de x car f est une immersion formelle en x, puis que g et h sont égaux sur un certain ouvert U de T contenant t car T est intègre et X séparé, et enfin comme U est dense dans T car celui-ci est irréductible, g et h sont égaux sur T car celui-ci est réduit et X séparé.

Démonstration. Les morphismes $\widehat{\mathcal{O}}_{Y,f(x)} \to \widehat{\mathcal{O}}_{X,x} \to \widehat{\mathcal{O}}_{T,t}$ déduits de $f \circ g$ et $f \circ h$ sont égaux, mais ce sont respectivement $\widehat{g}_t \circ \widehat{f}_x$ et $\widehat{h}_t \circ \widehat{f}_x$, or $\widehat{f}_x$ est surjective, donc $\widehat{g}_t = \widehat{h}_t$. L'anneau local $\mathcal{O}_{T,t}$ est intègre et noethérien donc $\mathcal{O}_{T,t} \to \widehat{\mathcal{O}}_{T,t}$ est injective par le théorème d'intersection de Krull ([AM94], Théorème 10.17 et son corollaire). Grâce au diagramme commutatif

$$\begin{array}{ccc} \widehat{\mathcal{O}}_{X,x} & \xrightarrow{\widehat{f_x}} & \widehat{\mathcal{O}}_{T,t} \\ \uparrow & & \uparrow \\ \mathcal{O}_{X,x} & \xrightarrow{f_x} & \mathcal{O}_{T,t} \end{array}$$

les homomorphismes $g_t, h_t : \mathcal{O}_{X,x} \to \mathcal{O}_{T,t}$ sont égaux. Ceci implique que g et h sont égaux grâce au lemme suivant.

Lemme I.2.14. *Soient T un schéma intègre et X un schéma séparé. Soient $g, h : T \to X$ deux morphismes tels que pour un certain $t \in T$, $g(t) = h(t) = x$ et les morphismes $g_t, h_t : \mathcal{O}_{X,x} \to \mathcal{O}_{T,t}$ sont égaux. Alors, $g = h$.*

Démonstration. Soit $V = \mathrm{Spec}\, A$ un ouvert affine de X contenant x, et $U = \mathrm{Spec}\, B$ un ouvert affine de T contenant t inclus dans $g^{-1}(V) \cap h^{-1}(V)$. Montrons que les morphismes $\mathrm{Spec}\, B \to \mathrm{Spec}\, A$ induits par g et h sont égaux. Notons ϕ et ψ les morphismes d'anneaux $A \to B$ associés. Par hypothèse, B est intègre et on a un certain idéal premier $\mathfrak{P}$ de B tel que $\phi^{-1}(\mathfrak{P}) = \psi^{-1}(\mathfrak{P}) = \mathfrak{p}$

et les morphismes induits $\phi_{\mathfrak{P}}, \psi_{\mathfrak{P}} : A_{\mathfrak{p}} \to B_{\mathfrak{P}}$ sont égaux. Les deux morphismes composés $A \to A_{\mathfrak{p}} \to B_{\mathfrak{P}}$ sont donc égaux, et comme B est intègre, $B \to B_{\mathfrak{P}}$ est injective, d'où $\phi = \psi$.

Les morphismes de schémas g et h coïncident donc sur l'ouvert U de T (comme morphisme de schémas et pas seulement comme applications). Le schéma T étant intègre donc irréductible, cet ouvert est dense dans T. Alors, comme g et h coïncident sur un ouvert dense de T, ils sont égaux car X est séparé et T est réduit ([Liu02], Proposition 3.3.11). □

□

Pour les applications qui nous intéressent, nous arriverons rarement à prouver directement que deux points ont même image par une immersion formelle. En général, on saura tout au plus montrer que la différence des deux images est de torsion dans une variété abélienne. C'est la raison pour laquelle on construit la propriété suivante pour les besoins de notre démonstration, qui est un raffinement de la proposition 3.1 de [Ell04] et le coeur de la méthode de Mazur.

Proposition I.2.15 (Propriété-clé de la méthode de Mazur)**.**

Soient K un corps de nombres et λ un idéal premier non nul de $\mathcal{O}_K$ de caractéristique résiduelle ℓ. On note $\mathcal{O}_\lambda$ le localisé de $\mathcal{O}_K$ en λ et $\mathbb{F}_\lambda$ le corps résiduel de $\mathcal{O}_K$ en λ.

Soient X une courbe algébrique définie sur $\mathbb{Q}$ qui admet un modèle propre $\mathcal{X}$ sur $\operatorname{Spec}\mathbb{Z}$, A une variété abélienne définie sur $\mathbb{Q}$ de modèle de Néron $\mathcal{A}$ sur $\operatorname{Spec}\mathbb{Z}$ et $f : X \to A$ un morphisme défini sur $\mathbb{Q}$. Celui-ci se prolonge canoniquement en un $\mathbb{Z}$-morphisme de schémas $f_{\mathbb{Z}} : \mathcal{X}^{\text{lisse}} \to \mathcal{A}$ par la propriété universelle du modèle de Néron de A. Soient enfin x et y deux points de $X(K)$ tels que :

- *Les points x et y ont même réduction modulo λ et celle-ci appartient à la partie lisse de $\mathcal{X}$.*
- *Le morphisme $f_{\mathbb{Z}}$ est une immersion formelle en $x_\lambda = y_\lambda$.*
- *Le point $f(y) - f(x)$ est $\mathbb{Q}$-rationnel et de torsion dans $A(\mathbb{Q})$.*

Alors, $x = y$ à moins que $\ell = 2$, et que $f(y) - f(x)$ soit un point d'ordre 2 de $A(\mathbb{Q})$ se réduisant en 0 modulo 2.

Ce résultat repose sur le lemme de spécialisation de Raynaud ([Maz77], Proposition 1.1), dont on donne une version simplifiée ici.

Proposition I.2.16 (lemme de spécialisation simplifié)**.**

Soit p un nombre premier. Soit K une extension de $\mathbb{Q}_p$ de ramification absolue $e < p - 1$ et de corps résiduel k. Alors, pour tout schéma en groupes fini plat G sur $\mathcal{O}_K$ et pour tout section $x \in G(\mathcal{O}_K)$, l'ordre de x est égal à l'ordre de sa spécialisation $x_k \in G(k)$. En particulier, la spécialisation est injective. Si $e = p - 1$, pour toute section $x \in G(\mathcal{O}_K)$, si la spécialisation x_k est nulle, x est nul ou exactement de p-torsion dans $G(\mathcal{O}_K)$ et il engendre alors une copie de μ_p dans G.

Au vu de cette proposition, le résultat précédent découle comme suit.

Démonstration. Plaçons nous tout d'abord dans le cas où $\ell > 2$. Soit $z = f(y) - f(x) \in A(\mathbb{Q})$. Cet élément est de torsion dans $A(\mathbb{Q})$ donc par le lemme de spécialisation de Raynaud, comme $e = 1 < \ell - 1$,

$$z_\lambda = z_\Lambda = f_{\mathbb{Z}}(x_\Lambda) - f_{\mathbb{Z}}(y_\Lambda) \in \mathcal{A}(\mathbb{F}_\ell)$$

est de même ordre de torsion que z. Or $x_\Lambda = y_\Lambda$ par hypothèse, donc $z = 0$. D'après la proposition I.2.13, $x = y$ car $f(x) = f(y)$ et f est une immersion formelle en $x_\Lambda = y_\Lambda$.

Pour $\ell = 2$, la fin de ce raisonnement (non dépendante du lemme de spécialisation) montre que $x = y$ pourvu que $z = 0$. Dans le cas contraire, z est un élément de torsion de $A(\mathbb{Q})$ qui s'annule modulo 2. D'après la deuxième partie de la proposition I.2.16, z est donc de 2-torsion. □

Remarques. Plusieurs points de la méthode sont à signaler ici :

- Dans la proposition I.2.15, on a besoin de savoir *a priori* que les réductions de x et y modulo λ sont dans la partie lisse de $\mathcal{X}$, pour pouvoir appliquer l'immersion formelle. En effet, la propriété d'extension de Néron ne prolonge le morphisme qu'à la partie lisse du modèle, ce qui nous oblige dans tous les cas, par un travail antérieur, à éliminer la possibilité que les points x qui nous concernent puissent se réduire dans la partie non lisse de $\mathcal{X}$. Nous verrons dans le paragraphe I.3.8 que ce travail, bien que fastidieux à première vue, nous gratifiera d'informations supplémentaires sur $f(y) - f(x)$.
- Prouver que $f(y) - f(x)$ est $\mathbb{Q}$-rationnel ne posera pas ici de gros problèmes si on a bien construit l'immersion formelle. La difficulté majeure est de prouver qu'il est de torsion. Le seul moyen dont on dispose est un argument de torsion « automatique », qui revient à prouver que $f(y) - f(x)$ est un point d'une sous-variété de A dont tous les points K-rationnels sont de torsion. A noter que l'utilisation d'une sous-variété de A plutôt qu'un quotient nous offre une plus grande souplesse que l'idée initiale de Mazur (notamment en ce qui concerne le cas $\ell = 2$). Cette idée technique est due à Merel notamment dans [Mer07].
- La principale différence avec la proposition 3.1 de [Ell04] est que celle-ci ne prenait pas en compte la possibilité d'avoir des points K-rationnels avec une image $\mathbb{Q}$-rationnelle à l'arrivée, ce qui sera notre cas (permettant ainsi de récupérer le cas $\ell = 3$). Elle ne permettait pas non plus de traiter le cas $\ell = 2$, fondamental pour la suite de notre étude.

I.2.2 Exemples remarquables d'immersions formelles

Soit p un nombre premier fixé. On travaillera ici avec $X = X_0(p)$ et $J_0(p)$ la jacobienne de X, toutes deux définies sur $\mathbb{Q}$, et avec le morphisme d'Albanese $\phi : X_0(p) \to J_0(p)$ qui envoie la pointe ∞ sur 0. Nous nous limiterons aux cas $p = 11$ ou $p > 13$ car ce sont les niveaux pour lesquels la courbe modulaire $X_0(p)$ est de genre non nul.

Par défaut, les résultats énoncés dans ce paragraphe ne seront pas démontrés, nous renvoyons au chapitre 3 de [Dar09] pour plus de détails.

L'algèbre de Hecke $\mathbb{T}$

On définit l'algèbre de Hecke $\mathbb{T}$ de la manière suivante : pour tout entier n non divisible par p, on a deux morphismes $\pi_1, \pi_2 : X_0(np)_{\mathbb{Q}} \to X_0(p)_{\mathbb{Q}}$, définis par

$$\pi_1(E, C) = (E, C_p) \quad \text{et} \quad \pi_2(E, C) = (E/C_n, C/C_n),$$

où C est un sous-groupe cyclique d'ordre np de E, C_p son unique sous-groupe de p-torsion, de cardinal p, et C_n son sous-groupe de n-torsion (remarquons que $\pi_2 = \pi_1 \circ w_n$ avec w_n l'involution d'Atkin-Lehner de $X_0(np)$ de degré n). Ces morphismes définissent une correspondance sur $X_0(p)_{\mathbb{Q}}$ définie sur $\mathbb{Q}$, qui à son tour donne un endomorphisme de $J_0(p)_{\mathbb{Q}}$ défini sur $\mathbb{Q}$, noté T_n. Lorsqu'on voit $J_0(p)$ comme le groupe de Picard de $X_0(p)$ par le théorème d'Abel-Jacobi, l'action de T_n sur les diviseurs est la suivante :

$$T_n.[E, C] = \sum_{\varphi : E \to E'} [E', C'],$$

la somme parcourant les isogénies $\varphi : E \to E'$ de noyau cyclique de cardinal n et $C' = \varphi(C)$. On note $\mathbb{T}$ le sous-anneau de $\mathrm{End}(J_0(p))$ engendré par les T_n et w_p (on renote $T_p = -w_p$). C'est une $\mathbb{Z}$-algèbre commutative qui est un $\mathbb{Z}$-module libre de rang fini, qu'on appelle *l'algèbre de Hecke de* $\Gamma_0(p)$. Elle agit naturellement sur les formes modulaires de $S_2(\Gamma_0(p), \mathbb{Z})$ (c'est-à-dire les formes modulaires de poids 2 pour $\Gamma_0(p)$ dont le développement en chaque pointe est à coefficients entiers). De plus, on peut expliciter cette action via le q-développement : pour tout $f = \sum_{n \geq 1} a_n q^n \in S_2(\Gamma_0(p), \mathbb{Z})$ et tout nombre premier ℓ,

$$T_\ell f = \begin{cases} \sum a_{\ell n} q^n + \ell \sum a_{\ell n} q^{\ell n} & \text{si } \ell \neq p, \\ \sum a_{\ell n} q^n & \text{si } \ell = p. \end{cases}$$

Grâce à ces formules, on démontre dans la proposition suivante que $\mathbb{T}$ et $S_2(\mathbb{Z}, \Gamma_0(p))$ sont en dualité parfaite.

Proposition I.2.17 (Dualité parfaite entre $\mathbb{T}$ et $S_2(\Gamma_0(p),\mathbb{Z})$)**.**

Soit $p = 11$ ou $p > 13$ un nombre premier. L'action de la $\mathbb{Z}$-algèbre $\mathbb{T}$ sur $S_2(\Gamma_0(p),\mathbb{Z})$ induit une dualité parfaite

$$\psi : \left| \begin{array}{ccc} \mathbb{T} \times S_2(\Gamma_0(p),\mathbb{Z}) & \longrightarrow & \mathbb{Z} \\ (T,f) & \longmapsto & a_1(T.f) \end{array} \right.$$

De plus, cette dualité est compatible à l'action de $\mathbb{T}$ au sens où pour tous les opérateurs $S,T \in \mathbb{T}$ et $f \in S_2(\Gamma_0(p),\mathbb{Z})$,

$$\psi(ST,f) = \psi(S,T.f) = \psi(T,S.f).$$

On en déduit un isomorphisme canonique de $\mathbb{T}$-modules

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}) \longrightarrow S_2(\Gamma_0(p),\mathbb{Z})$$

avec $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})$ muni de sa structure canonique de $\mathbb{T}$-module $(T.\psi)(T') = \psi(TT')$ pour tout $T,T' \in \mathbb{T}$ et $\psi \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})$.

Démonstration. Il suffit de prouver que le morphisme induit $\varphi : S_2(\Gamma_0(p),\mathbb{Z}) \to \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})$ défini par $\varphi(f)(T) = \psi(T,f)$ est un isomorphisme.

Il est injectif car si $\psi(T,f) = 0$ pour tout $T \in \mathbb{T}$, en particulier $a_n(f) = a_1(T_n f) = 0$ pour tout $n \in \mathbb{N}$ donc $T = 0$. Son image est de même rang que celui de $\mathbb{T}$ (donc de conoyau fini) car l'application $\mathbb{T} \to \mathrm{Hom}(S_2(\Gamma_0(p),\mathbb{Z}),\mathbb{Z})$ induite par ψ est elle aussi injective : en effet, si $\psi(T,f) = 0$ pour tout $f \in S_2(\Gamma_0(p),\mathbb{Z})$, en particulier $\psi(T,T_n \cdot f) = 0$ pour tout $n \in \mathbb{N}$ et tout $f \in S_2(\Gamma_0(p),\mathbb{Z})$ donc $T \cdot f = 0$ pour tout $f \in S_2(\Gamma_0(p),\mathbb{Z})$ par le même argument que précédemment, donc $T = 0$ en tant qu'endomorphisme de $S_2(\Gamma_0(p),\mathbb{Z})$, or $S_2(\Gamma_0(p),\mathbb{Z})$, en tant que $\mathbb{T}$-module, est l'espace cotangent de $J_0(p)$ le long de la section nulle (voir le principe de q-développement ci-dessous), donc $T = 0$ en tant qu'opérateur sur cet espace cotangent, donc en tant qu'endomorphisme de $J_0(p)$, soit $T = 0$. Le morphisme φ est donc injectif de conoyau fini, ainsi pour $L \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})$, il existe $m \in \mathbb{N}^*$ et $f \in S_2(\Gamma_0(p),\mathbb{Z})$ unique tels que $mL(T) = \psi(T,f)$ pour tout $T \in \mathbb{T}$. Alors, pour tout $T \in \mathbb{T}$, $a_1(T/mf) = L(T) \in \mathbb{Z}$ donc pour tout $n \in \mathbb{N}^*$, $a_n(f) \in m\mathbb{Z}$, ce qui prouve que $f/m \in S_2(\Gamma_0(p),\mathbb{Z})$ par définition de cet espace, et alors $L = \varphi(f/m)$, ce qui prouve la surjectivité de φ. □

Le principe de q-développement

On note $J_0(p)_{\mathbb{Z}}$ le modèle de Néron de la variété abélienne $J_0(p)_{\mathbb{Q}}$ sur $\mathbb{Z}$.

On note $\phi_{\mathbb{Z}} : X_0(p)^{\mathrm{lisse}}_{\mathbb{Z}} \to J_0(p)_{\mathbb{Z}}$ l'extension par propriété de Néron du morphisme d'Albanese $\phi : X_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ défini dans la section précédente.

L'algèbre de Hecke $\mathbb{T}$ agit sur $J_0(p)_{\mathbb{Z}}$ par la propriété universelle du modèle de Néron, et donc en particulier sur l'espace cotangent en la section nulle. D'après Mazur ([Maz78], §2.(e) ou plus précisément [Edi84], début de la preuve du théorème 3.2 pour le résultat sur $\mathbb{Z}$), on a un isomorphisme de $\mathbb{T}$-modules

$$\mathrm{Cot}_0(J_0(p)_{\mathbb{Z}}) \longrightarrow S_2(\Gamma_0(p),\mathbb{Z})$$

où $\mathrm{Cot}_0(J_0(p)_{\mathbb{Z}})$ est le $\mathbb{Z}$-module cotangent à la section nulle $0_{\mathbb{Z}}$ sur $J_0(p)_{\mathbb{Z}}$, et selon cet isomorphisme, le diagramme

$$\begin{array}{ccc} \mathrm{Cot}_0(J_0(p)_{\mathbb{Z}}) & \xrightarrow{\cong} & S_2(\Gamma_0(p),\mathbb{Z}) \\ {\scriptstyle \phi^*_{\mathbb{Z}}}\downarrow & & \downarrow \\ \mathrm{Cot}_\infty(X_0(p)_{\mathbb{Z}}) & \xrightarrow{\cong} & \mathbb{Z} \end{array} \qquad \begin{array}{c} \sum_{n\geq 1} a_n q^n \\ \top \\ \downarrow \\ a_1 \end{array}$$

où $\phi^*_{\mathbb{Z}}$ est le morphisme induit par $\phi_{\mathbb{Z}}$ sur les espaces cotangents, commute au signe près.

Comme $S_2(\Gamma_0(p),\mathbb{Z})$ s'identifie à $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})$ (proposition I.2.17), on obtient le diagramme commutatif au signe près suivant.

$$\begin{array}{ccc} \mathrm{Cot}_0(J_0(p)_{\mathbb{Z}}) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}) \\ {\scriptstyle \phi_{\mathbb{Z}}^*}\downarrow & & \downarrow \\ \mathrm{Cot}_\infty(X_0(p)_{\mathbb{Z}}) & \xrightarrow{\cong} & \mathbb{Z} \end{array} \qquad \begin{array}{c} \psi \\ \top \\ \downarrow \\ \psi(1_{\mathbb{T}}) \end{array}$$

Si on choisit $T \in \mathbb{T}$, c'est un endomorphisme de $J_0(p)_{\mathbb{Z}}$, et son action sur $\mathrm{Cot}_0(J_0(p)_{\mathbb{Z}})$ est donnée par le diagramme commutatif

$$\begin{array}{ccc} \mathrm{Cot}_0(J_0(p)_{\mathbb{Z}}) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}) \\ {\scriptstyle T^*}\downarrow & & \downarrow \\ \mathrm{Cot}_0(J_0(p)_{\mathbb{Z}}) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}) \end{array} \qquad \begin{array}{c} \psi \\ \top \\ \downarrow \\ \psi \circ T \end{array}$$

car via l'isomorphisme $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}) \to S_2(\Gamma_0(p),\mathbb{Z})$, on peut voir l'application de T sur une forme modulaire comme la multiplication à gauche par T. Nous avons maintenant assez pour trouver une famille d'exemples d'immersions formelles, dont on fera grand usage par la suite.

Proposition I.2.18 (Immersions formelles sur $X_0(p)$)**.**

Soit $p = 11$ ou $p > 13$ un nombre premier. Soit $\phi : X_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ le morphisme d'Albanese envoyant la pointe ∞ sur 0 et $\phi_{\mathbb{Z}} : X_0(p)_{\mathbb{Z}}^{\mathrm{lisse}} \to J_0(p)_{\mathbb{Z}}$ son extension par la propriété universelle de Néron.

Pour tout $T \in \mathbb{T}$ et tout ℓ premier, la composition $T \circ \phi_{\mathbb{Z}}$ est une immersion formelle en ∞_ℓ si et seulement si $T \notin \ell\mathbb{T}$.

Démonstration. Soient ℓ un nombre premier et $T \in \mathbb{T}$, notons $\phi_T = T \circ \phi_{\mathbb{Z}}$. Remarquons que ϕ_T envoie ∞_ℓ sur 0_ℓ dans la jacobienne, et les corps résiduels sont tous les deux $\mathbb{F}_\ell$. L'équivalence porte donc, par la caractérisation des immersions formelles (proposition I.2.10), exclusivement sur les espaces cotangents en ∞_ℓ et 0_ℓ. Par définition des espaces cotangents de sections, l'application cotangente ϕ_T en ∞_ℓ est exactement $\phi_T^* \otimes \mathbb{F}_\ell$. Via les diagrammes commutatifs ci-dessus, cette application est au signe près

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})/\ell.\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}) & \longrightarrow & \mathbb{Z}/\ell\mathbb{Z} \\ \overline{\psi} & \longmapsto & \overline{\psi(T)} \end{array}$$

L'espace d'arrivée étant de dimension 1 sur $\mathbb{F}_\ell$, l'application cotangente est surjective si et seulement si elle est non nulle. Supposons que pour tout $\psi \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z})$, $\psi(T) \in \ell\mathbb{Z}$. Alors, l'application $\psi \mapsto \psi(T)$ appartient à $\ell.\mathrm{Hom}(\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T},\mathbb{Z}),\mathbb{Z})$ qui par bidualité n'est autre que $\ell\mathbb{T} \subset \mathbb{T}$, donc $T \in \ell\mathbb{T}$. Réciproquement, il est clair que si $T \in \ell\mathbb{T}$, l'application cotangente en ∞_ℓ est nulle. □

Ces immersions formelles seront l'outil adéquat pour nos démonstrations futures. Nous étant chargés de trouver les immersions formelles, cherchons maintenant les variétés abéliennes d'arrivée convenables, et la torsion rationnelle de la jacobienne.

I.3 Quotient d'Eisenstein et groupe des composantes de $J_0(p)$

Cette section regroupe les résultats principaux sur le quotient d'Eisenstein (préféré au quotient d'enroulement dans les cas Borel et normalisateur de Cartan déployé pour son caractère plus explicite), et sur le groupe des composantes, dont nous nous servirons particulièrement pour la bonne réduction modulo 2 dans les cas « Borel » et « normalisateur de Cartan déployé ».

I.3.1 Idéal d'Eisenstein et quotient d'Eisenstein

Le quotient d'Eisenstein a été défini, étudié et utilisé par Mazur [Maz77] dans le but de fournir une variété abélienne quotient de $J_0(p)$ qui dispose d'un nombre fini de points $\mathbb{Q}$-rationnels (c'est-à-dire qu'il est de rang zéro sur $\mathbb{Q}$).

Définition I.3.1 (Idéal et quotient d'Eisenstein)**.**
L'*idéal d'Eisenstein* $\mathcal{I}$ de $\mathbb{T}$ est l'idéal

$$\mathcal{I} = \langle 1 + \ell - T_\ell, 1 + w_p, l \in \mathcal{P}_p \rangle.$$

où $\mathcal{P}_p$ désigne l'ensemble des nombres premiers différents de p. On note également

$$\gamma_{\mathcal{I}} = \ker(\mathbb{T} \to \varprojlim \mathbb{T}/\mathcal{I}^n) = \bigcap_n \mathcal{I}^n.$$

Le *quotient d'Eisenstein de* $J_0(p)$ est le quotient de la jacobienne $J_0(p)$ par la sous-variété abélienne engendrée par $\gamma_{\mathcal{I}}.J_0(p)$. On note ce quotient $\widetilde{J}(p)$, et il est défini sur $\mathbb{Q}$.

La dénomination « idéal d'Eisenstein » provient du fait qu'on retrouve les valeurs propres des séries d'Eisenstein pour les opérateurs de Hecke dans cette définition.

L'article de Mazur est pour une grande part consacré à l'étude de ce quotient, qui est le premier exemple non trivial de quotient de $J_0(p)$ de rang zéro sur $\mathbb{Q}$. Plus précisément, sont démontrés dans cet article les résultats suivants.

Proposition I.3.2 (Propriétés fondamentales du quotient d'Eisenstein)**.**
Soit $p = 11$ *ou* $p > 13$ *un nombre premier. Soit* $n = \mathrm{num}(\frac{p-1}{12})$. *Soit* C *le sous-groupe de* $J_0(p)(\mathbb{Q})$ *engendré par* $\mathrm{cl}((0) - (\infty))$. *Alors :*

(*a*) *La torsion rationnelle de* $J_0(p)$ *est exactement* C, *et c'est un sous-groupe cyclique d'ordre* n *([Maz77], Théorème 1.2 p. 142 et proposition 11.1 p. 98).*

(*b*) *La projection* $J_0(p) \to \widetilde{J}(p)$ *est définie sur* $\mathbb{Q}$ *et met en bijection* $C = J_0(p)(\mathbb{Q})_{\mathrm{tors}}$ *avec* $\widetilde{J}(p)(\mathbb{Q})$, *qui est donc un groupe cyclique d'ordre* n *([Maz77], Corollaire 1.4 p.143).*

(*c*) *L'idéal d'Eisenstein n'est autre que le noyau de l'application* $T \mapsto T.\mathrm{cl}((0) - (\infty))$ *de* $\mathbb{T}$ *dans* C. *En conséquence, l'évaluation en* $\mathrm{cl}((0) - (\infty))$ *induit un isomorphisme de* $\mathbb{Z}$*-modules* $\mathbb{T}/\mathcal{I} \to \mathbb{Z}/n\mathbb{Z}$ *([Maz77], Proposition 11.1).*

Remarque I.3.1. Le groupe C est appelé *groupe cuspidal* car c'est exactement le groupe engendré par les diviseurs de degré zéro sur $X_0(p)$ à support dans les pointes. On a donc ici sa structure exacte, mais il faut signaler que le théorème de Drinfeld-Manin suffit à démontrer qu'il est de torsion. La deuxième partie du (a) peut donc être vue comme une explicitation de Drinfeld-Manin dans notre cas.

On sait maintenant envoyer des points de $X_0(p)(\mathbb{Q})$ sur des points de $J_0(p)(\mathbb{Q})_{\mathrm{tors}}$ de la manière suivante. Soit T un opérateur de Hecke tel que $T.\gamma_{\mathcal{I}} = 0$. Alors, T est nul sur la sous-variété abélienne engendrée par $\gamma_{\mathcal{I}}.J_0(p)$ donc en tant qu'endomorphisme de $J_0(p)$, l'opérateur T se factorise par $\widetilde{J}(p)$. En particulier, il envoie tout point rationnel de $J_0(p)$ sur un point rationnel *de torsion* de $J_0(p)$, c'est-à-dire un élément du groupe cuspidal C d'après la proposition I.3.2. Pour $\phi : X_0(p) \to J_0(p)$ le morphisme d'Albanese de point-base ∞, le morphisme $T \circ \phi$ envoie donc bien $X_0(p)(\mathbb{Q})$ dans $C = J_0(p)(\mathbb{Q})_{\mathrm{tors}}$. Cette construction nous permettra d'utiliser à bon escient la proposition I.2.15 dans les sections suivantes.

Il sera important dans les sections I.4 et I.5 d'annuler $\gamma_{\mathcal{I}}$ par des éléments de $\mathbb{T}$ dont on maîtrise la congruence modulo $\mathcal{I}$, d'où le lemme suivant.

Lemme I.3.3. *Pour tout nombre premier* ℓ, *il existe un opérateur de Hecke* $T \in \mathbb{T}\backslash\ell\mathbb{T}$ *tel que* $T = 1 \mod \mathcal{I}$ *et* $T \cdot \gamma_{\mathcal{I}} = 0$. *De plus, pour un tel* $T \in \mathbb{T}$, *on a* $T \cdot (1 + w_p) = 0$.

Démonstration. Pour commencer, un tel $T \in \mathbb{T}$ annule automatiquement $(1+w_p)$ car le quotient d'Eisenstein est un quotient de la partie négative de $J_0(p)$ ([Maz77], Proposition 17.10). Montrons maintenant qu'il en existe. Comme l'algèbre $\mathbb{T}$ est noethérienne, $\gamma_{\mathcal{I}}$ est un $\mathbb{T}$-module de type fini et il existe donc $T \in \mathbb{T}$ congru à 1 modulo $\mathcal{I}$ tel que $T \cdot \gamma_{\mathcal{I}} = 0$, par le théorème d'intersection de Krull ([AM94], Théorème 10.17). Pour tout nombre premier ℓ, on peut même choisir un tel annulateur n'appartenant pas à $\ell\mathbb{T}$: si ℓ divise n, c'est automatiquement le cas, sinon, on peut choisir un nombre premier ℓ' différent de ℓ mais congru à ℓ modulo n. Alors, si $T \in \ell\mathbb{T}$, il existe $k \in \mathbb{N}_{>0}$ tel que $T \in \ell^k\mathbb{T} - \ell^{k+1}\mathbb{T}$. Alors, l'opérateur $T' = (\ell'/\ell)^k.T \notin \ell\mathbb{T}$ annule $\gamma_{\mathcal{I}}$ et est congru à 1 modulo $\mathcal{I}$. □

I.3.2 Fibre en p de $X_0(p)$ et groupe des composantes

L'idée motrice de cette partie est que non seulement on sait envoyer un point de $X_0(p)(\mathbb{Q})$ sur un point de $J_0(p)(\mathbb{Q})_{\text{tors}}$ (paragraphe précédent), mais on est en fait capable de contrôler les images possibles dans le groupe cuspidal. Ceci sera crucial pour la réduction modulo 2, où on a besoin de savoir que cette image n'est pas de 2-torsion pour éviter l'exception de la proposition I.2.15.

Pour acquérir cette information supplémentaire, nous allons étudier la fibre géométrique en p du modèle minimal régulier de $X_0(p)$ sur $\mathbb{Z}$, et via un théorème de Raynaud en déduire quelles sont les possibles réductions d'éléments de $X_0(p)$ dans le groupe des composantes de la jacobienne. Les opérateurs de Hecke agissant naturellement sur ce groupe, on n'aura plus qu'à en déduire les images possibles, sachant que le groupe cuspidal se réduit injectivement dans ce groupe des composantes. Tout ceci est détaillé dans les propositions suivantes. Cette section est pour l'essentiel une traduction de la section 2.3 de [LF].

La base théorique de l'étude du groupe des composantes est le résultat suivant ([BLR90], Théorème 9.6.1).

Proposition I.3.4. *Soit R un anneau de valuation discrète de corps des fractions K et de corps résiduel k supposé parfait.*

Soit X une courbe propre, plate et régulière sur R telle que X_K est géométriquement irréductible.

Soit J_R le modèle de Néron de la jacobienne J de X sur R. On note $\mathcal{C}$ l'ensemble des composantes irréductibles de $X_{\overline{k}}$, toutes supposées lisses, et Φ le groupe des composantes connexes de $(J_R)_{\overline{k}}$. Alors, on a un isomorphisme $\ker\beta/\operatorname{im}\alpha \cong \Phi$, *avec*

$$\alpha : \left| \begin{array}{ccc} \bigoplus_{C\in\mathcal{C}} \mathbb{Z}.[C] & \longrightarrow & \bigoplus_{C\in\mathcal{C}} \mathbb{Z}.[C] \\ [C] & \longmapsto & \sum_{C'\in\mathcal{C}} (C\cdot C')[C'] \end{array} \right. \qquad \beta : \left| \begin{array}{ccc} \bigoplus_{C\in\mathcal{C}} \mathbb{Z}.[C] & \longrightarrow & \mathbb{Z} \\ \sum_{C\in\mathcal{C}} \lambda_C[C] & \longmapsto & \sum_{C\in\mathcal{C}} \lambda_C \end{array} \right. ,$$

où $(C\cdot C')$ est le nombre d'intersection de C et C' en tant que diviseurs de Cartier. Via cet isomorphisme, pour tout $P \in J(K)$, si D est un élément de J_R tel que la restriction de D à J_K est exactement P, alors P se spécialise dans $(J_R)_{\overline{k}}$ dans la composante correspondant à

$$\sum_{C\in\mathcal{C}} (D\cdot C)[C].$$

On note maintenant R un anneau de valuation discrète complet et de caractéristique nulle, K son corps des fractions et k son corps résiduel, de caractéristique p et supposé parfait. On note également π une uniformisante de R, v la valuation discrète normalisée par $v(\pi) = 1$ et $e = v(p)$ la ramification absolue de R.

La proposition suivante (tirée de la section 3 de l'Annexe de [BD97]) généralise le théorème 1.1 de l'Annexe de [Maz77]. Précisons que ce résultat supposait R de corps résiduel algébriquement clos, mais la formation des modèles minimal et de Néron commute au changement de base étale, on peut donc l'appliquer ici. On pourra également reprendre les résultats pour K un corps de nombres sur $\mathbb{Q}$ et un idéal premier $\mathfrak{P}$ de $\mathcal{O}_K$ au-dessus de p.

Proposition I.3.5 (Structure de la fibre spéciale de $X_0(p)_R$)**.**

(a) *Le schéma $X_0(p)_R$ est lisse sur R en-dehors des points singuliers de sa fibre spéciale.*

(b) *La fibre $X_0(p)_{\overline{k}}$ est constituée de deux copies de $\mathbb{P}^1(\overline{k})$ qui s'intersectent transversalement en les modules de courbes elliptiques supersingulières. La première copie, notée Z, paramètre les courbes elliptiques munies de leur isogénie de Frobenius, et la seconde, notée Z', paramètre celles qui sont munies de leur isogénie Verschiebung.*

(c) *Soit s un point de $X_0(p)_{\overline{k}}$ correspondant à une paire (E, C_p), avec E une courbe elliptique sur $\overline{k}$ et C_p une p-isogénie de E. On appelle épaisseur de s l'entier $k_s = |\mathrm{Aut}(E, C_p)|/2$. L'anneau local complété de $X_0(p)_R$ en s est isomorphe à $R[[X,Y]]/(XY - \pi^{ek_s})$, en particulier le schéma $X_0(p)_R$ est non régulier en s si et seulement si $ek_s > 1$.*

Si $p \neq 2, 3$, alors $k_s > 1$ implique que $j(E) = 0$, $k_s = 3$ et $p = -1 \mod 3$, ou $j(E) = 1728$, $k_s = 2$ et $p = -1 \mod 4$.

(d) *La fibre $(\widetilde{X_0(p)_R})_{\overline{k}}$ du modèle régulier minimal $\widetilde{X_0(p)}_R$ de $X_0(p)$ sur R est obtenue par éclatement de $(X_0(p))_{\overline{k}}$ en chaque point non régulier s de cette fibre en une chaîne de $ek_s - 1$ droites projectives $\mathbb{P}^1_{\overline{k}}$. En tant que diviseurs de Cartier, ces droites projectives ont pour auto-intersection -2.*

Une première application de ce résultat à $J_0(p)_{\mathbb{Z}}$ donne la proposition suivante ([Maz77], Théorème 10 et Annexe).

Proposition I.3.6. *Soient $p = 11$ ou $p > 13$ un nombre premier et $n = \mathrm{num}\left(\frac{p-1}{12}\right)$. On note $\phi : X_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ le morphisme d'Albanese de point-base ∞, et $J_0(p)_{\mathbb{Z}}$ le modèle de Néron de $J_0(p)$ sur $\mathbb{Z}$.*

(a) *Le groupe des composantes Φ de $(J_0(p)_{\mathbb{Z}})_{\overline{\mathbb{F}_p}}$ est cyclique d'ordre n, engendré par $\overline{Z}$. On l'identifie à $(\mathbb{Z}/n\mathbb{Z})$ avec ce choix de générateur.*

(b) *La réduction modulo p met en bijection C et Φ en envoyant $\mathrm{cl}((0) - (\infty))$ sur $\overline{Z}$.*

(c) *Pour tout point $P = (E, C_p) \in Y_0(p)(\mathbb{Q})$, on note $\rho(P) \in \mathbb{Z}/n\mathbb{Z}$ l'image de P par la composition*

$$X_0(p)(\mathbb{Q}) \to J_0(p)(\mathbb{Q}) \to J_0(p)_{\mathbb{Z}}(\overline{\mathbb{F}_p}) \to \Phi \cong \mathbb{Z}/n\mathbb{Z}.$$

Alors, pour tout point $P \in Y_0(p)(\mathbb{Q})$:

- *Si E a réduction potentiellement ordinaire ou multiplicative, alors $\rho(\phi(P)) = 0$ si C_p définit une isogénie séparable modulo p, et $\rho(\phi(P)) = 1$ sinon.*
- *Sinon, soit $p = -1 \mod 4$, $j(E) = 0 \mod p$ et alors $\rho(\phi(P)) = 1/2$ soit $p = -1 \mod 3$, soit $j(E) = 1728 \mod p$ et alors $\rho(\phi(P)) = 1/3$ ou $2/3$.*

L'idée du (c) de cette proposition est que $\rho(\phi(P))$ ne peut pas être $n/2$, c'est-à-dire l'unique point de torsion non trivial de $\mathbb{Z}/n\mathbb{Z}$, à moins que n (donc p) soit petit. L'article de Momose et Shimura [MS02] donne une justification supplémentaire (et plus détaillée) du fait que la réduction modulo p d'une courbe elliptique associée à un point de $X_0(p)(\mathbb{Q})$ ne peut être supersingulière, sauf peut-être si son j-invariant est égal à 0 ou 1728 modulo p. En particulier, il explique que cela est impossible si la courbe elliptique en question est semi-stable sur $\mathbb{Q}$ par un argument général sur les groupes formels à un paramètre sur $\mathbb{Z}_p$.

La proposition précédente permet de retrouver la rétraction de $J_0(p)(\mathbb{Q})$ dans le groupe cuspidal identifiée par Mazur ([Maz77], p. 99).

Corollaire I.3.1. *Soit $p = 11$ ou $p > 13$ un nombre premier. Soit $n = \mathrm{num}(\frac{p-1}{12})$. Soit C le groupe cuspidal de $J_0(p)$ (qui est rationnel), $C_{\mathbb{Z}}$ son adhérence schématique dans le modèle de Néron $J_0(p)_{\mathbb{Z}}$ et $\overline{C}$ la spécialisation en p de $C_{\mathbb{Z}}$ dans $J_0(p) \otimes_{\mathbb{F}_p} \overline{\mathbb{F}_p}$.*

(a) *La réduction modulo p induit une bijection entre C et $\overline{C}$.*

(b) *Le groupe $\overline{C}$ est un système de représentants des composantes irréductibles de $J_0(p)_{\overline{\mathbb{F}_p}}$.*

(c) *On a donc une rétraction $\rho : J_0(p)(\mathbb{Q}) \to C$ compatible avec la réduction modulo p et avec l'action de l'algèbre de Hecke.*

Remarque I.3.2. L'action de l'algèbre de Hecke sur le groupe des composantes Φ se déduit de l'action de $\mathbb{T}$ sur C par spécialisation, et selon cette action T est la multiplication par k dans $\mathbb{Z}/n\mathbb{Z}$ si $T = k \mod \mathcal{I}$. Comme ρ se projette en l'identité dans le groupe des composantes, elle est compatible à l'action de l'algèbre de Hecke car la réduction modulo p l'est.

Pour établir en toute généralité la structure du groupe des composantes, on a besoin de notations supplémentaires. Tout d'abord, on peut supposer que $e > 1$ car le cas $e = 1$ est fait dans la proposition I.3.6, de sorte que tout module s de courbe elliptiques supersingulière sur $\overline{k}$ est un point non régulier de $X_0(p)_R$.

Pour fournir une interprétation plus visuelle de la preuve, on définit (comme dans l'Annexe de [BD97]) le graphe dual $\widetilde{G}$ associé à $\widetilde{(X_0(p)_R)_{\overline{k}}}$: ses sommets sont les composantes irréductibles de $\widetilde{(X_0(p)_R)_{\overline{k}}}$ et les arêtes entre deux composantes distinctes correspondent à leurs points d'intersection. La proposition I.3.4 transforme donc notre problème de groupe des composantes en un problème sur $\widetilde{G}$, c'est-à-dire calculer le groupe abélien Φ dont les générateurs sont les sommets de $\widetilde{G}$ et les relations sont données par l'image de l'opérateur laplacien sur $\widetilde{G}$.

Remarquons de plus que les relations entre composantes correspondent exactement à la loi de Kirchoff appliquée en chaque point du graphe d'intersection (dont les sommets sont les composantes irréductibles et les arêtes les intersections entre ces composantes). On notera en conséquence les relations suivant le point en lequel on applique cette loi par la suite.

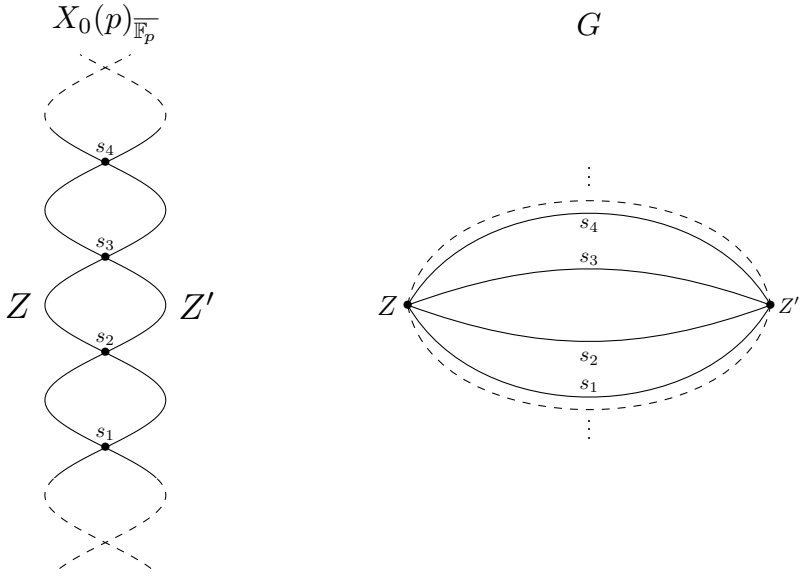
Soit $\mathcal{S}$ (resp. $\mathcal{S}'$) l'ensemble de cardinal S (resp. S') des points supersinguliers de $X_0(p)_{\overline{k}}$ (resp. supersinguliers de j-invariant différent de 0 et 1728 modulo p). On note aussi $I = 1$ (resp. $R = 0$) si la courbe elliptique de j-invariant 1728 (resp. 0) est supersingulière dans k, et 0 sinon. On a alors les deux formules

$$S = S' + I + R \quad \text{et} \quad S' + \frac{I}{2} + \frac{R}{3} = \frac{p-1}{12}, \tag{I.4}$$

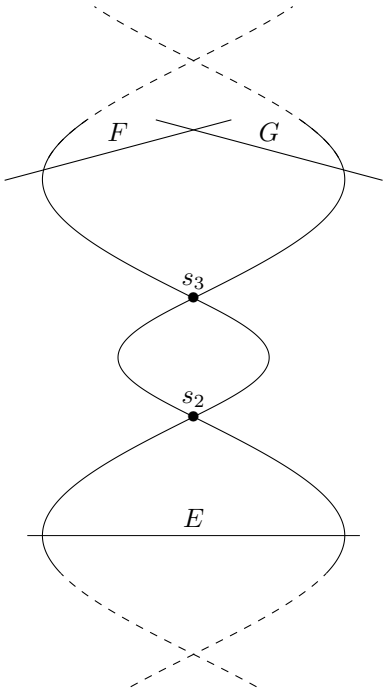
d'après le théorème V.4.1 (c) de [Sil09].

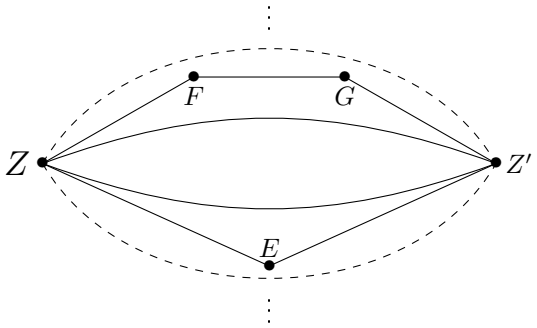
Pour tout $s \in \mathcal{S}$, on appelle $\mathcal{C}_s$ le chemin de longueur ek_s dans $\widetilde{G}$ entre les points Z et Z' associé au point $s \in \mathcal{S}$. Dans le cas du j-invariant 1728 (resp. 0), on va aussi l'appeler $\mathcal{E}$ (resp. $\mathcal{G}$). Par convention, on ordonne les points du chemin $\mathcal{C}_s$ de la manière suivante : on note $C_{s,0} = Z'$, $C_{s,1} = C_s$ l'unique point de $\mathcal{C}_s$ relié à Z', $C_{s,2}$ l'unique point de $\mathcal{C}_s$ relié à C_s non encore nommé, et ainsi de suite jusqu'à $C_{s,ek_s} = Z$. Si s est de j-invariant 1728 (resp. 0), on note $E = \mathcal{C}_{s,1}$ (resp. $G = \mathcal{C}_{s,1}$) pour retrouver les notations de l'Annexe de [Maz77] (où $F = \mathcal{C}_{s,2}$ avec $j(s) = 0$).

Les dessins suivants illustrent le comportement des composantes irréductibles en p de $X_0(p)$ et de son modèle minimal régulier (à droite, on peut voir le graphe dual).



$\widetilde{X_0(p)_{\overline{\mathbb{F}_p}}}$ $(p \equiv 1 \mod 12)$ $\widetilde{G}$





Pour toute composante irréductible C, on note

$$\overline{C} = [C] - [Z'],$$

comme Z' est la composante contenant la réduction de la pointe ∞, notre choix de point-base pour le morphisme d'Albanese. Le lemme suivant sert à simplifier la présentation de Φ.

Lemme I.3.7. *Soit $\mathcal{C}_s$ le chemin de longueur ek_s entre Z et Z' associé à s. Dans le groupe Φ, pour tout $i \in \{0, \cdots, ek_s\}$,*

$$\overline{C_{s,i}} = i\overline{C_{s,1}} = i\overline{C_s}.$$

En particulier,

$$\overline{Z} = ek_s\overline{C_s}.$$

Démonstration. C'est vrai par définition de nos systèmes d'indices pour $i = 0$ et 1. De plus, pour tout $i \in \{1, \cdots, m-1\}$, la relation donnée par l'opérateur laplacien sur $\widetilde{G}$ en en $C_{s,i}$ est exactement

$$-2\overline{C_{s,i}} + \overline{C_{s,i-1}} + \overline{C_{s,i+1}} = 0.$$

Le résultat en découle par récurrence double sur i. □

Grâce à ce lemme, le groupe Φ est en fait le groupe abélien de générateurs $\overline{Z}$ et les $\overline{C_s}$, $s \in \mathcal{S}$, avec les relations

$$\begin{array}{rcll} -S\overline{Z} + (2e-1)I\overline{E} + (3e-1)R\overline{G} + \sum_{s\in\mathcal{S}'}(e-1)\overline{C_s} & = & 0 & (Z) \\ I\overline{E} + R\overline{G} + \sum_{s\in\mathcal{S}'}\overline{C_s} & = & 0 & (Z') \\ \overline{Z} & = & e\overline{C_s}, & (C_s), s \in \mathcal{S}' \end{array}$$

avec les relations supplémentaires

$$\begin{array}{rcll} \overline{Z} & = & 2e\overline{E} & (E) \\ \overline{Z} & = & 3e\overline{G} & (G) \end{array}$$

quand les composantes mentionnées existent. En ajoutant à la relation (Z) la relation $-(e-1)(Z')$, on obtient une nouvelle relation (Z)

$$-S\overline{Z} + eI\overline{E} + 2eR\overline{G} = 0 \qquad (Z).$$

qu'on utilise à la place de la précédente. On peut maintenant donner la structure de Φ.

Proposition I.3.8 (Groupe des composantes dans le cas général)**.**
Soit $p = 11$ ou $p > 13$ un nombre premier et $n = \mathrm{num}\left(\frac{p-1}{12}\right)$. Avec les notations et définitions de toute cette section, on note $J_0(p)_R$ le modèle de Néron de $J_0(p)$ sur R.

(a) Le groupe des composantes irréductibles Φ de $(J_0(p)_R)_{\overline{k}}$ est non-canoniquement de la forme

$$\Phi \cong (\mathbb{Z}/ne\mathbb{Z}) \times (\mathbb{Z}/e\mathbb{Z})^{S-2}.$$

(b) Le groupe cuspidal C de $J_0(p)(\mathbb{Q})$ se réduit de manière injective dans Φ, avec $\overline{Z}$ la réduction de $\mathrm{cl}([0]-[\infty])$. En conséquence, le groupe engendré par $\overline{Z}$ dans Φ est de cardinal n, et on l'identifie à $(\mathbb{Z}/n\mathbb{Z})$ grâce à ce choix de générateur.

(c) On a une suite exacte de $\mathbb{Z}/e\mathbb{Z}$-modules

$$0 \longrightarrow \mathbb{Z}/e\mathbb{Z} \xrightarrow{\Delta} (\mathbb{Z}/e\mathbb{Z})^S \xrightarrow{\alpha} \Phi/\langle\overline{Z}\rangle \longrightarrow 0$$

avec $\Delta : \lambda \mapsto \lambda \cdot \sum_{s\in S}[C_s]$ et $\alpha : \sum_{s\in S}\lambda_s[C_s] \mapsto \sum_{s\in S}\lambda_s\overline{C_s}$. En particulier, $e\cdot\Phi = \langle\overline{Z}\rangle$, et on a

$$\begin{array}{rcll}
\forall s \in \mathcal{S}', \forall i \in \{1,\cdots,e-1\}, e\cdot\overline{C_{s,i}} & = & i & \\
\forall i \in \{1,\cdots,2e-1\}, e\cdot\overline{E_i} & = & i/2 & \text{si } p=-1 \mod 4 \\
\forall i \in \{1,\cdots,3e-1\}, e\cdot\overline{G_i} & = & i/3 & \text{si } p=-1 \mod 3 \\
\sum_{s\in\mathcal{S}}\overline{C_s} & = & 0 & \text{dans } \Phi
\end{array}$$

Démonstration. Le calcul est similaire dans chacun des quatre cas de congruences modulo 12 (qui détermine l'existence ou non des composantes exceptionnelles E et G), nous allons le faire seulement pour $p = 11 \mod 12$ qui est légèrement plus technique que les autres. On remplace $\overline{Z}$ par $2e\overline{E}$ avec la relation (E), et pour tout $s \in \mathcal{S}'$, on fait le changement de variables

$$\overline{C_s'} := \overline{C_s} - 2\overline{E} \quad \text{and} \quad \overline{G'} := \overline{G} + (2S-3)\overline{E}.$$

Les relations deviennent alors

$$\begin{array}{rcll}
e(6S-7)\overline{E} + 2e\overline{G'} & = & 0 & (Z)' \\
\overline{G'} + \sum_{s\in\mathcal{S}'}\overline{C_s'} & = & 0 & (Z')' \\
e\overline{C_s'} & = & 0 & (C_s)', s \in \mathcal{S}'
\end{array}$$

Comme $p = 11 \mod 12$, on a $6S - 7 = 6S' + 5 = (p-1)/2 = n$ par la formule (I.4), et $(Z)'$ est équivalente à $en\overline{E} = 0$ avec les autres relations. En conséquence, $\overline{G'}$ est dans le groupe engendré par les autres générateurs, et les relations $(C_s)', s \in \mathcal{S}'$ et $(Z)'$ sont diagonales, ce qui nous donne l'isomorphisme annoncé du (a), par lequel

$$\begin{array}{rcl}
\overline{Z} & = & (2e, 0, \cdots, 0) \\
\overline{E_i} & = & (i, 0\cdots, 0) \\
\overline{G_i} & = & (-(2S'+1)i, -i, \cdots, -i) \\
\overline{C_{s,i}} & = & (2i, 0, \cdots, 0, i, 0, \cdots, 0) \quad (s \in \mathcal{S}')
\end{array}$$

Ceci nous donne directement le (b) et le (c). □

Nous allons nous servir de cette proposition dans les cas des sous-groupes de Borel et de Cartan déployé, car elle permet d'imposer des contraintes supplémentaires et ainsi de prouver la bonne réduction potentielle en caractéristique 2.

I.4 Le cas des sous-groupes de Borel

Soient p un nombre premier fixé et d un entier sans facteur carré non divisible par p. Le schéma $X_0(dp)_{\mathbb{Z}}$ (voir section I.1.4) est lisse sur $\mathbb{Z}$ en-dehors des points supersinguliers en caractéristique divisant dp, en particulier toutes ses pointes se réduisent dans la partie lisse modulo tout nombre premier ℓ. Si r est le nombre de facteurs premiers de d, $X_0(dp)_{\mathbb{Q}}$ a 2^{r+1} pointes sur lesquelles le groupe des involutions d'Atkin-Lehner agit transitivement.

On note $\pi_{dp,p} : X_0(dp)_{\mathbb{Q}} \to X_0(p)_{\mathbb{Q}}$ et $\pi_{dp,d} : X_0(dp)_{\mathbb{Q}} \to X_0(d)_{\mathbb{Q}}$ les morphismes d'oubli respectivement de la d-structure et de la p-structure. Comme les niveaux sont premiers entre eux, les pointes de $X_0(dp)_{\mathbb{Q}}$ correspondent via ces morphismes aux couples de pointes de $X_0(d)_{\mathbb{Q}}$ et $X_0(p)_{\mathbb{Q}}$. Pour cette correspondance, observons que l'involution d'Atkin-Lehner w_d ne change pas la composante des pointes en $X_0(p)_{\mathbb{Q}}$, inversement w_p ne change pas la composante des pointes en $X_0(d)_{\mathbb{Q}}$ mais permute 0 et ∞ dans $X_0(p)_{\mathbb{Q}}$. On note ∞^{dp} la pointe infinie usuelle de $X_0(dp)_{\mathbb{Q}}$ et ∞ celle de $X_0(p)_{\mathbb{Q}}$.

Définition I.4.1. Soit $\phi : X_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ le morphisme d'Albanese envoyant ∞ sur 0. On définit alors $g : X_0(dp)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ par $g = \phi \circ \pi_{dp,p} + \phi \circ \pi_{dp,p} \circ w_d$. Fonctoriellement, on a

$$g(E, C_d, C_p) = \mathrm{cl}([E, C_p] + [E/C_d, C_p + C_d/C_d] - 2[\infty]).$$

Pour tout $T \in \mathbb{T}$, on note $g_T = T \circ g$.

Cette famille de morphismes contient les candidats pour les immersions formelles.

Proposition I.4.2 (Immersions formelles, cas Borel)**.**
Soit $p = 11$ ou $p > 13$ un nombre premier. Soit ℓ un nombre premier éventuellement égal à p. Pour tout $T \in \mathbb{T}$, le morphisme $(g_T)_{\mathbb{Z}} : X_0(dp)^{\mathrm{lisse}}_{\mathbb{Z}} \to J_0(p)_{\mathbb{Z}}$ étend g_T par propriété de Néron est une immersion formelle en ∞^{dp}_ℓ si et seulement si $T \notin \ell\mathbb{T}$.

Démonstration. Fixons ℓ un nombre premier et $T \in \mathbb{T}$. Notons $g'_T = T \circ \phi \circ \pi_{dp,p}$ de sorte que $g_T = g'_T + g'_T \circ w_d$. Remarquons tout d'abord que $g'_T(\infty^{dp}) = g'_T \circ w_d(\infty^{dp}) = 0$ car w_d permute les pointes au-dessus de ∞. Les corps résiduels de ∞^{dp}_ℓ et 0_ℓ sont tous les deux $\mathbb{F}_\ell$, il reste donc à vérifier le critère des morphismes cotangents d'après la caractérisation des immersions formelles (proposition I.2.10). L'application cotangente de $(\pi_{dp,p})_{\mathbb{Z}} : X_0(dp)_{\mathbb{Z}} \to X_0(p)_{\mathbb{Z}}$ en la section ∞^{dp} est l'identité sur $\mathbb{Z}$, donc l'application cotangente de g'_T en ∞^{dp}_ℓ est surjective si et seulement si $T \notin \ell\mathbb{T}$ d'après la proposition I.2.18. Ensuite, $w_d(\infty^{dp})$ est une pointe de $X_0(dp)_{\mathbb{Q}}$ différente de ∞^{dp}, et la projection $\pi_{dp,p}$ est donc ramifiée en cette pointe (la projection naturelle de surfaces de Riemann $X_0(d)_{\mathbb{C}} \to X(1)_{\mathbb{C}}$ est ramifiée en toute pointe sauf la pointe infinie). Ainsi, l'application cotangente de $g'_T \circ w_d$ en la section ∞^{dp} est nulle, en particulier nulle en ∞^{dp}_ℓ. L'application cotangente de g_T en ∞^{dp}_ℓ est donc celle de g'_T, ainsi g_T est une immersion formelle en ∞^{dp}_ℓ si et seulement si T n'appartient pas à $\ell\mathbb{T}$. □

Grâce à cette preuve, nous pouvons prouver notre théorème de bonne réduction. Le lemme suivant est important pour le cas $\ell = 2$, mettons-le à part pour plus de clarté.

Lemme I.4.3 (Images dans le groupe cuspidal)**.**
Soit $p = 11$ ou $p > 13$ un nombre premier et $n = \mathrm{num}(p - 1/12)$. Soit K un corps quadratique.
Soient E/K une $\mathbb{Q}$-courbe de degré d sans facteur carré telle que $\mathbb{P}\overline{\rho}_{E,p}$ est réductible, et P le point de $X_0(dp)(K)$ correspondant. Alors, $g(P)$ est un point de $J_0(p)(\mathbb{Q})$, et via la rétraction $\rho : J_0(p)(\mathbb{Q}) \to C \cong \mathbb{Z}/n\mathbb{Z}$ du corollaire I.3.1, on a le tableau des valeurs possibles de $\rho(g(P))$ suivantes, avec e l'indice de ramification de p dans K :

	$p = 1 \mod 12$	$p = 5 \mod 12$	$p = 7 \mod 12$	$p = 11 \mod 12$
$e = 1$	$0, 2$	$0, 1, 2, \frac{2}{3}, \frac{4}{3}$	$0, 1, 2$	$0, 1, 2, \frac{2}{3}, \frac{4}{3}$
$e = 2$	$0, 1, 2$	$0, 1, 2, \frac{1}{3}, \frac{2}{3}, \frac{4}{3}, \frac{5}{3}$	$0, 1, 2, \frac{1}{2}, \frac{3}{2}$	$0, 1, 2, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}$

En particulier, cette image ne peut être le point de 2-torsion non trivial de $\mathbb{Z}/n\mathbb{Z}$ lorsque $p \geq 11$ et $p \neq 13, 17, 41$.

Démonstration. Pour σ l'automorphisme non trivial de K, on a

$$^{\sigma}g(P) = \mathrm{cl}([^{\sigma}\phi(P)] + [^{\sigma}\phi(w_d.P)] - 2[^{\sigma}\infty]) = \mathrm{cl}([\phi(^{\sigma}P)] + [\phi(w_d.^{\sigma}P)] - 2[^{\sigma}\infty])$$

ce qui mène à

$$^{\sigma}g(P) = \mathrm{cl}([\phi(w_d.P)] + [\phi(P)] - 2[\infty]) = g(P)$$

par construction du point P (sous-section I.1.17), donc $g(P)$ est bien rationnel.

Ensuite, le groupe de torsion de $J_0(p)(\mathbb{Q})$ est le groupe cuspidal C, cyclique de cardinal n et la réduction modulo p induit une bijection entre celui-ci et le groupe des composantes du modèle de Néron de $J_0(p)$ sur $\mathbb{Z}_p$ (proposition I.3.2). On peut donc lire $\rho(g(P))$ dans le groupe des composantes. Pour cela, il faut la lire dans le groupe des composantes Φ du modèle de Néron de $J_0(p)$ sur $(\mathcal{O}_K)_{\mathfrak{P}}$, avec $\mathfrak{P}$ un idéal premier de $\mathcal{O}_K$ au-dessus de p.

Remarquons tout d'abord que les points P et $w_d.P$ représentent des courbes elliptiques isogènes de degré premier à p sur $X_0(p)$. En conséquence, leur réduction semi-stable modulo $\mathfrak{P}$ (ordinaire, supersingulière ou multiplicative) est de même nature.Ainsi, les deux composantes obtenues l'une par réduction de $\phi \circ \pi(P)$, l'autre par réduction de $\phi \circ \pi(w_d \cdot P)$ sont soit toutes les deux Z, soit toutes les deux Z', soit toutes les deux des composantes supersingulières (appartenant aux mêmes chemins $\mathcal{E}$ ou $\mathcal{G}$, comme deux points de ces différents chemins ne peuvent être isogènes). Cette remarque faite, on utilise la proposition I.3.8 pour calculer les possibilités.

• Si p est non ramifié dans $\mathcal{O}_K$, on est dans le cas étale de la proposition, et les valeurs possibles de $g(P)$ sont donc $2\overline{Z'} = 0$, $2\overline{Z} = 2$, $2\overline{E} = 1$ quand E existe, $2\overline{G} = 2/3$, $2\overline{F} = 4/3$, et $\overline{F} + \overline{G} = 1$ quand F, G existent. Cela prouve la première ligne du tableau.

• Si p est ramifié dans $\mathcal{O}_K$, on est dans le cas $e = 2$ de la proposition, et les valeurs possibles de $g(P)$ sont alors $2\overline{Z'} = 0$, $2\overline{Z} = 2$, $2\overline{E_1} = 1/2$, $2\overline{E_2} = 1$, $2\overline{E_3} = 3/2$, $\overline{E_1} + \overline{E_3} = 1$, $2\overline{G_1} = 1/3$, $2\overline{G_2} = 2/3$, $2\overline{G_3} = 1$, $2\overline{G_4} = 4/3$, $2\overline{G_5} = 5/3$, et les sommes entre ces différents G_i qui mènent aux mêmes possibilités.

Pour l'application du résultat, notons que $\mathbb{Z}/n\mathbb{Z}$ admet un point de 2-torsion non trivial si et seulement si 2 divise n, soit si et seulement si $p = 1 \mod 8$, ce qui élimine la moitié des cas, et qu'alors ce point est l'image de $n/2$. Supposons que c'est le cas pour notre courbe E. Si $p = 1 \mod 12$, on a $n = (p-1)/12$ qui divise $1, 2$ ou 4, d'où $p = 13$.

Si $p = 5 \mod 12$, on a $n = (p-1)/4$ qui divise $2, 4, 8$ ou 10, ce qui donne $p = 17$ ou $p = 41$.

Une fois ces cas exclus, on est donc certain que $\rho(g(P))$ n'est pas le point de 2-torsion non trivial de $\mathbb{Z}/n\mathbb{Z}$. □

Remarquons tout d'abord qu'on peut reprouver le cas rationnel avec notre lemme, en guise d'échauffement pour le cas des $\mathbb{Q}$-courbes.

Proposition I.4.4 (Bonne réduction, cas rationnel)**.** *Soit E une courbe elliptique définie sur $\mathbb{Q}$ et p un nombre premier tel que $\rho_{E,p}$ est à valeurs dans un sous-groupe de Borel de* $\mathrm{GL}(E[p])$*. Alors, pour $p = 11$ ou $p > 13$, la courbe elliptique E a potentiellement bonne réduction en tout nombre premier.*

Démonstration. Soit $P = (E, C_p) \in X_0(p)(\mathbb{Q})$ associé à E par hypothèse. On reprend les notations du paragraphe I.2.2. Soit ℓ un nombre premier (qui peut être égal à p), supposons par l'absurde que E a réduction potentiellement multiplicative modulo ℓ. Alors, P se réduit en une pointe modulo ℓ, et quitte à utiliser une involution d'Atkin-Lehner (qui ne change pas la nature de la réduction de E), on peut supposer que $P_\ell = \infty_\ell$. Soit $T \in \mathbb{T}\backslash\ell\mathbb{T}$ annulant l'idéal $\gamma_{\mathcal{I}}$ définissant le quotient d'Eisenstein (lemme I.3.3). Le morphisme $T : J_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ est nul sur $\gamma_{\mathcal{I}}.J_0(p)$ et se factorise donc par la projection $J_0(p) \to \tilde{J}(p)$ en un morphisme $\mathbb{Q}$-rationnel $\tilde{J}(p) \to J_0(p)$. En particulier, il envoie tout point $\mathbb{Q}$-rationnel de $J_0(p)$ sur l'image d'un point $\mathbb{Q}$-rationnel de $\tilde{J}(p)$. Ceux-ci étant de torsion, $\phi_T(P)$ est un point de torsion de $J_0(p)(\mathbb{Q})$, et on peut donc appliquer la proposition I.2.15 à $x = P$, $y = \infty$, $X = X_0(p)$ et $A = J_0(p)$. Dans le cas où $\ell > 2$, on obtient $P = \infty$, contradiction, donc E a potentiellement bonne réduction modulo ℓ. Dans le cas où $\ell = 2$, le raisonnement tient toujours excepté lorsque $\phi_T(P)$ est l'unique point d'ordre exactement 2 du groupe cuspidal, et en supposant de plus que T est congru à 1 modulo $\mathcal{I}$ (lemme I.3.3), cela ne peut arriver d'après le

cas étale du lemme I.4.3 que pour $p = 2, 3, 5, 13$, qui sont des cas déja exclus, donc pour $\ell = 2$, on obtient encore $P = \infty$, ce qui contredit notre hypothèse de départ et conclut donc la preuve par l'absurde. □

Théorème I.3 (Bonne réduction, cas des $\mathbb{Q}$-courbes)**.** *Soient K un corps quadratique, et E/K une $\mathbb{Q}$-courbe de degré d sans facteur carré. Si pour $p \geq 11$, $p \neq 13, 17, 41$ premier ne divisant pas d, la représentation $\mathbb{P}\overline{\rho}_{E,p}$ est réductible, E a potentiellement bonne réduction en tout idéal premier de $\mathcal{O}_K$.*

Remarque I.4.1. Grâce au lemme I.4.3 et à la proposition I.2.15, ce théorème améliore la proposition 3.2 de [Ell04] en traitant aussi les idéaux premiers de $\mathcal{O}_K$ au-dessus de 2 et 3. Il est également valide pour les courbes elliptiques définies sur $\mathbb{Q}$. En particulier, on sait que les points rationnels non cuspidaux non CM de $X_0(p)$ ont potentiellement bonne réduction partout, c'est-à-dire que leur j-invariant est entier (les cas $p = 17$ et $p = 41$ n'ont pas lieu d'être dans le cas étale du lemme 5.2, on retrouve donc la condition classique $p = 11$ ou $p > 13$). Dans l'article original de Mazur, le cas $\ell = 2$ était exclu, mais il était inutile pour la fin de la méthode employée par Mazur lui-même.

Démonstration. Soit P le point de $X_0(dp)(K)$ associé à E par la preuve de la proposition I.1.17. D'après le lemme I.4.3, le point $g(P)$ est $\mathbb{Q}$-rationnel.

Soit maintenant λ un idéal premier de $\mathcal{O}_K$ de caractéristique ℓ, notons $\mathcal{O}_\lambda$ le localisé de $\mathcal{O}_K$ en λ. Supposons par l'absurde que P a réduction potentiellement multiplicative en λ. Alors, P_λ est la réduction d'une pointe de $X_0(dp)$ modulo λ. Le groupe des involutions d'Atkin-Lehner agissant transitivement et $\mathbb{Q}$-rationnellement sur les pointes de $X_0(dp)$, on peut et on va supposer que $P_\lambda = \infty_\ell$. En particulier, P_ℓ appartient à la partie lisse de $X_0(dp)_{\mathbb{Z}}$, et ce même si ℓ divise dp.

Soit $T \in \mathbb{T}\backslash\ell\mathbb{T}$ annulant l'idéal $\gamma_{\mathcal{I}}$ définissant le quotient d'Eisenstein (lemme I.3.3). Le morphisme $T : J_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ est nul sur $\gamma_{\mathcal{I}}.J_0(p)$ et se factorise donc par la projection $J_0(p) \to \tilde{J}(p)$ en un morphisme $\mathbb{Q}$-rationnel $\tilde{J}(p) \to J_0(p)$. En particulier, il envoie tout point $\mathbb{Q}$-rationnel de $J_0(p)$ sur l'image d'un point $\mathbb{Q}$-rationnel de $\tilde{J}(p)$. Ceux-ci étant de torsion, $g_T(P)$ est un point de torsion de $J_0(p)(\mathbb{Q})$. On peut donc appliquer la proposition I.2.15 à $x = P$, $y = \infty^{dp}$, $X = X_0(dp)$, $A = J_0(p)$. Dans le cas où $\ell > 2$, on obtient $P = \infty^{dp}$, contradiction, donc E a potentiellement bonne réduction modulo ℓ.

Dans le cas où $\ell = 2$, si $P \neq \infty^{dp}$, $g_T(P)$ est un point de 2-torsion non-trivial de $J_0(p)(\mathbb{Q})$ grâce à la proposition I.2.15. On peut ici de plus choisir T congru à 1 modulo $\mathcal{I}$ (lemme I.3.3), de sorte que la composante de $g_T(P)$ est la même que la composante de $g(P)$ dans le groupe des composantes de $J_0(p)\otimes\overline{\mathbb{F}_p}$ car la rétraction ρ est $\mathbb{T}$-équivariante (corollaire I.3.1). D'après le lemme I.4.3, il est impossible que ce point soit de 2-torsion non trivial sous notre hypothèse $p \neq 13, 17, 41$, donc E a potentiellement bonne réduction en tout idéal premier de $\mathcal{O}_K$. □

I.5 Le cas des normalisateurs de sous-groupes de Cartan déployés

Soient $p = 11$ ou $p > 13$ un nombre premier fixé et $d > 1$ un entier sans facteur carré non divisible par p.

On note $X_0^{\mathrm{sC}}(d;p)_{\mathbb{Z}} = X_0(d)_{\mathbb{Z}} \times_{X(1)} X_{\mathrm{sp.Car.}}(p)_{\mathbb{Z}}$ le schéma de modules grossier paramétrant les quadruplets (E, C_d, A_p, B_p) avec E une courbe elliptique, C_d un sous-groupe cyclique d'ordre d et A_p, B_p deux sous-groupes cycliques d'ordre p distincts de E. Ce schéma est propre sur $\operatorname{Spec}\mathbb{Z}$, de fibre générique notée $X_0^{\mathrm{sC}}(d;p)_{\mathbb{Q}}$. Il est naturellement muni d'une involution

$$w : (E, C_d, A_p, B_p) \mapsto (E, C_d, B_p, A_p)$$

de sorte que le quotient $X_0^{\mathrm{sC}}(d;p)_{\mathbb{Z}}/\langle w\rangle$ n'est autre que le schéma $X_0^{\mathrm{s}}(d;p)_{\mathbb{Z}}$ défini dans le paragraphe I.1.4. Nous allons donc procéder en définissant nos morphismes sur $X_0^{\mathrm{sC}}(d;p)_{\mathbb{Z}}$ avant de les faire passer au quotient.

Comme p et d sont premiers entre eux, les morphismes d'oubli de structure $X_0^{\mathrm{sC}}(d;p) \to X_0(d)$ et $X_0^{\mathrm{sC}}(d;p) \to X_{\mathrm{sp.Car.}}(p)$ mettent en bijection les pointes de $X_0^{\mathrm{sC}}(d;p)$ et les couples de pointes

de $X_0(d)$ et de $X_{\text{sp.Car.}}(p)$. On notera donc (c, c') la pointe de $X_0^{\text{sC}}(d;p)$ au-dessus de la pointe c de $X_0(d)$ et c' de $X_{\text{sp.Car.}}(p)$. De même, les morphismes $X_0^{\text{s}}(d;p) \to X_0(d)$ et $X_0^{\text{s}}(d;p) \to X_{\text{split}}(p)$ mettent en bijection les pointes de $X_0^{\text{s}}(d;p)$ et les couples de pointes de $X_0(d)$ et de $X_{\text{split}}(p)$, on emploiera la même notation.

Le schéma $X_0^{\text{sC}}(d;p)_{\mathbb{Z}}$ est lisse en-dehors des caractéristiques divisant dp. Pour les caractéristiques ℓ divisant d, ses points singuliers sont les points supersinguliers en caractéristique ℓ, en particulier il est lisse en toutes les réductions de pointes modulo ℓ. Pour la caractéristique p, $X_0^{\text{sC}}(d;p)_{\overline{\mathbb{F}_p}}$ est constituée (Partie 1 de [Mom84]) de trois composantes irréductibles qui hors de leurs intersections paramètrent les quadruplets (E, C_d, A_p, B_p) tels que respectivement :

- Le schéma en groupes A_p est étale-localement isomorphe à μ_p sur $\overline{\mathbb{F}_p}$, en particulier l'isogénie associée est inséparable (composante Z).
- Le schéma en groupes B_p est étale-localement isomorphe à μ_p sur $\overline{\mathbb{F}_p}$ (composante Z').
- Ni A_p ni B_p ne sont étale-localement isomorphes à μ_p (composante W).

Les pointes de $X_0^{\text{sC}}(d;p)_{\overline{\mathbb{F}_p}}$ au-dessus de ∞ dans $X_{\text{sp.Car.}}(p)_{\mathbb{F}_p}$ appartiennent à la composante Z, les pointes au-dessus de 0 dans $X_{\text{sp.Car.}}(p)$ à la composante Z', et les autres pointes de $X_0^{\text{sC}}(d;p)_{\overline{\mathbb{F}_p}}$ à la composante W, qui est de multiplicité $p-1$. Ces dernières sont donc des points non lisses de $X_0^{\text{sC}}(d;p)$. Plus généralement, un point de $X_0^{\text{sC}}(d;p)_{\overline{\mathbb{F}_p}}$ est singulier si et seulement s'il appartient à la composante W (c'est en particulier le cas quand il est supersingulier).

Par passage au quotient dans $X_0^{\text{s}}(d;p)$, les composantes Z et Z' s'identifient dans la fibre en p de $X_0^{\text{s}}(d;p)$, qui est donc constituée de cette image (notée Z_0) et de celle de W (renotée W). Les pointes n'appartenant pas à W sont donc les pointes au-dessus de la pointe ∞ de $X_{\text{split}}(p)$, les autres sont des points singuliers.

La proposition suivante condense les propriétés particulières d'une courbe elliptique dans le cas « normalisateur de Cartan déployé », d'après le lemme 1.3 de [Mom84].

Proposition I.5.1. *Soient E une courbe elliptique définie sur un corps de nombres K. Soit $p > 2[K:\mathbb{Q}]+1$ tels que $\rho_{E,p}$ est à valeurs dans le normalisateur d'un sous-groupe de Cartan déployé de $\mathrm{GL}(E[p])$. Soit $P = (E, \{A_p, B_p\}$ le point de $X_{\text{split}}(p)(K)$ associé, et K' une extension de degré 2 de K telle que A_p et B_p sont stables par $\mathrm{Gal}(\overline{\mathbb{Q}}/K')$. Alors, pour tout idéal premier $\mathfrak{P}$ de $\mathcal{O}_K$ au-dessus de p :*

(a) La réduction semi-stable de E modulo $\mathfrak{P}$ n'est pas supersingulière.

(b) La réduction modulo $\mathfrak{P}$ de $(E, \{A_p, B_p\})$ dans $X_{\text{split}}(p)_{\mathbb{F}_p}$ n'appartient pas à la composante W.

(c) Pour tout idéal premier $\mathfrak{P}'$ de $\mathcal{O}_{K'}$ au-dessus de $\mathfrak{P}$, les réductions modulo $\mathfrak{P}'$ de (E, A_p) et $(E/B_p, E[p]/B_p)$ appartiennent à la même composante de $X_0(p)_{\mathbb{F}_p}$.

Démonstration. Soit $p \geq 5$ un nombre premier quelconque et $\mathfrak{P}'$ un idéal premier de $\mathcal{O}' := \mathcal{O}_{K'}$ au-dessus de p, de ramification absolue e'.

Les $\mathrm{Gal}(\overline{\mathbb{Q}}/K')$-modules A_p et B_p définissent des schémas en groupes sur K' notés $(A_p)_{K'}$ et $(B_p)_{K'}$, et de la décomposition $E[p] = A_p \oplus B_p$ on déduit que $E[p]$ définit lui aussi un schéma en groupes sur K' tel que

$$(E[p])_{K'} \cong (A_p)_{K'} \oplus (B_p)_{K'}. \tag{I.5}$$

Par passage à l'adhérence schématique dans le modèle de Néron de E sur $\mathcal{O}'$, on obtient des schémas en groupes $(E[p])_{\mathcal{O}'}, (A_p)_{\mathcal{O}'}$ et $(B_p)_{\mathcal{O}'}$ sur $\mathcal{O}'$ prolongeant les schémas en groupes sur K'. Comme $e' \leq 2[K:\mathbb{Q}] < p-1$, d'après le lemme de spécialisation de Raynaud ([Maz78], proposition 1.1), nous avons donc

$$E_{\mathcal{O}'}[p] \cong (E[p])_{\mathcal{O}'} \cong (A_p)_{\mathcal{O}'} \oplus (B_p)_{\mathcal{O}'}.$$

En conséquence, comme d'après ce même lemme de spécialisation, les schémas en groupes (de rang p) $(A_p)_{\mathcal{O}'}$ et $(B_p)_{\mathcal{O}'}$ sont constants ou isomorphes à μ_p, le schéma en groupes $E_{\mathcal{O}'}[p]$ ne peut pas contenir le schéma en groupes α_p, ce qui interdit à E d'avoir réduction potentiellement supersingulière, soit le (a).

Ensuite, le schéma en groupes $(E[p])_{\mathcal{O}'}$ ne peut pas être étale car ses points géométriques dans la fibre en p sont au plus au nombre de p, et il est de rang p^2. L'un des schémas de droite ne peut donc pas être étale, il est alors isomorphe à μ_p, ce qui prouve le (b).

Enfin, la décomposition (I.5) induit un isomorphisme entre $(A_p)_{K'}$ et $(E[p]/B_p)_{K'}$. Comme $e' \leq 2[K : \mathbb{Q}] < p-1$, encore une fois d'après le lemme de spécialisation de Raynaud, cet isomorphisme s'étend en un isomorphisme $(A_p)_{\mathcal{O}'} \cong (E[p]/B_p)_{\mathcal{O}'}$. En particulier, ces deux schémas en groupes sont simultanément étales ou radiciels. Comme la composante à laquelle appartient (E, A_p) dans $X_0(p)_{\overline{\mathbb{F}_p}}$ est déterminée par la nature étale ou non de A_p, les deux points (E, A_p) et $(E/B_p, E[p]/B_p)$ appartiennent à la même composante de $X_0(p)_{\overline{\mathbb{F}_p}}$.

□

Remarque I.5.1. Dans le cas Borel ce n'était pas un problème (les pointes étant dans le domaine lisse), mais ici il faut savoir que la réduction des points qui nous intéressent est bien dans le domaine lisse pour pouvoir utiliser les immersions formelles, c'est entre autres pourquoi la proposition précédente est indispensable pour la réduction modulo p.

Voyons maintenant quelle famille d'immersions formelles sera adaptée à notre étude. Rappelons que $\phi : X_0(p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ est le morphisme d'Albanese envoyant ∞ sur 0 et que $X_0^{\mathrm{sC}}(d;p)_{\mathbb{Z}}$ est munie de deux involutions : une involution w_d qui descend en l'involution de Fricke sur $X_0(d)$ par la projection et une involution w qui permute les sous-groupe cycliques d'ordre p.

Définition I.5.2. Soit p un nombre premier. On note $\pi : X_0^{\mathrm{sC}}(d;p)_{\mathbb{Q}} \to X_0(p)_{\mathbb{Q}}$ le morphisme d'oubli qui envoie (E, C_d, A_p, B_p) sur (E, A_p) et w_p l'involution d'Atkin-Lehner de $X_0(p)$. On définit l'application $h : X_0^{\mathrm{sC}}(d;p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$ par

$$h = \mathrm{cl}(\phi \circ \pi - \phi \circ w_p \circ (\pi \circ w) + \phi \circ \pi \circ w_d - \phi \circ w_p \circ (\pi \circ w \circ w_d)).$$

Cette expression se traduit fonctoriellement par

$$\begin{aligned} h(E, C_d, A_p, B_p) &= \mathrm{cl}([E, A_p] - [E/B_p, E[p]/B_p] \\ &+ [E/C_d, (A_p + C_d)/C_d] - [E/(B_p + C_d), (E[p] + C_d)/(B_p + C_d)]). \end{aligned}$$

Par construction, on a $h \circ w_d = h$ et $h \circ w = -w_p \circ h$ où w_p désigne par abus de notation l'endomorphisme de $J_0(p)_{\mathbb{Q}}$ correspondant à l'involution d'Atkin-Lehner $w_p : X_0(p) \to X_0(p)$. On note pour tout $T \in \mathbb{T}$, $h_T = T \circ h$. Pour tout T tel que $T \cdot (1 + w_p) = 0$, on a $h_T \circ w = h_T$ donc h_T se factorise via la projection $X_0^{\mathrm{sC}}(d;p)_{\mathbb{Q}} \to X_0^{\mathrm{s}}(d;p)_{\mathbb{Q}}$ en un morphisme $\mathbb{Q}$-rationnel $h_T^+ : X_0^{\mathrm{s}}(d;p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$. Par abus de notation, on désignera par les mêmes lettres les extensions de ces morphismes au-dessus de $\mathbb{Z}$ par propriété de Néron.

Remarque I.5.2. Les deux premiers termes de h constituent le bon candidat h' pour les courbes elliptiques sur $\mathbb{Q}$, voir par exemple [BP11a]. Comme dans le cas des sous-groupes de Borel, le passage aux $\mathbb{Q}$-courbes consiste à considérer $h' + h' \circ w_d$.

Proposition I.5.3 (Immersions formelles, cas déployé)**.**
Pour tout nombre premier ℓ et tout $T \in \mathbb{T}$ (resp. tel que $T(1 + w_p) = 0$), le morphisme h_T (resp. h_T^+) est une immersion formelle en la pointe $(\infty, \infty)_\ell$ de $X_0^{\mathrm{sC}}(d;p)_{\mathbb{F}_\ell}$ (resp. $X_0^{\mathrm{s}}(d;p)_{\mathbb{F}_\ell}$) si et seulement si $T \notin \ell\mathbb{T}$.

Démonstration. Soit $\psi : X_0^{\mathrm{sC}}(d;p)_{\mathbb{Q}} \to X_0(dp)_{\mathbb{Q}}$ le morphisme d'« oubli de B_p », c'est-à-dire qui envoie (E, C_d, A_p, B_p) sur (E, C_d, A_p). Avec les définitions de la section I.4, on a $\pi = \pi_{dp,p} \circ \psi$ et $h = g \circ \psi + g \circ w_p \circ \psi \circ w$. Comme $\psi((\infty, \infty)) = \infty^{dp} = w_p \circ \psi \circ w(\infty, \infty)$, l'application cotangente de $h_{\mathbb{Z}}$ le long de la section $(\infty, \infty)_{\mathbb{Z}}$ est la somme des applications cotangentes de $(g \circ \psi)_{\mathbb{Z}}$ et $(g \circ w_p \circ \psi \circ w)_{\mathbb{Z}}$. De plus, ψ est ramifiée de degré p en $(\infty, \infty)_{\mathbb{Z}}$ (comme on peut le constater immédiatement sur les surfaces de Riemann), donc $g \circ \psi$ l'est. L'application cotangente de $h_{\mathbb{Z}}$ est donc exactement celle de $(g \circ w_p \circ \psi \circ w)_{\mathbb{Z}}$. De plus, on a $g \circ w_p = w_p \circ g + 2\,\mathrm{cl}([0] - [\infty^p])$, ainsi $(h_t)_{\mathbb{Z}}$ est une immersion formelle en $(\infty, \infty)_{\mathbb{F}_\ell}$ si et seulement si $(g_t \circ \psi \circ w)_{\mathbb{Z}}$ l'est, car w_p est un automorphisme de $J_0(p)_{\mathbb{Z}}$. Maintenant, l'application cotangente de $(\psi \circ w)_{\mathbb{Z}} : X_0^{\mathrm{sC}}(d;p)_{\mathbb{Z}} \to X_0(p)_{\mathbb{Z}}$ le long de la section $(\infty, \infty)_{\mathbb{Z}}$ est un isomorphisme ([Mom84], Preuve de la proposition 2.5), d'où le résultat pour $(h_t)_{\mathbb{Z}}$ par la proposition I.4.2. Il implique directement la propriété pour $(h_t)^+_{\mathbb{Z}}$ car (∞, ∞) n'est pas un point fixe de w. □

La proposition suivante montre, autant dans son énoncé que dans sa preuve, en quoi ce cas-ci est différent du cas Borel.

Proposition I.5.4 (Annulation automatique par h_T^+, cas déployé)**.**
Soit E une $\mathbb{Q}$-courbe de degré d sur le corps quadratique K et $p = 11$ ou $p > 13$ un nombre premier tels que $\mathbb{P}\overline{\rho}_{E,p}$ est à valeurs dans le normalisateur d'un sous-groupe de Cartan déployé de $\mathrm{GL}(E[p])$*. Soit P le point de $X_0^{\mathrm{s}}(d;p)(K)$ associé à E. Pour tout $T \in \mathbb{T}$ qui annule l'idéal $\gamma_{\mathcal{I}}$, on a automatiquement $T(1+w_p) = 0$ et $h_T^+(P) = 0$.*

Démonstration. D'après la proposition 17.10 de [Maz77], pour un tel $T \in \mathbb{T}$, on a $T(1+w_p) = 0$, donc h_T se factorise bien en $h_T^+ : X_0^{\mathrm{s}}(d;p)_{\mathbb{Q}} \to J_0(p)_{\mathbb{Q}}$. Le point $h_T^+(P)$ est K-rationnel car h_T^+ est défini sur $\mathbb{Q}$ et P est K-rationnel, de plus pour σ l'automorphisme non trivial de $\mathrm{Gal}(K/\mathbb{Q})$,

$$^{\sigma}h_T^+(P) = h_T^+(^{\sigma}P) = h_T^+(w_d.P) = h_T^+(P)$$

car $^{\sigma}P = w_d.P$ (proposition I.1.17) et $h_T \circ w_d = h_T$ par construction, donc $h_T^+ \circ w_d = h_T^+$. Ainsi, $h_T^+(P)$ est un point $\mathbb{Q}$-rationnel de $T.J_0(p)$. Cette sous-variété abélienne de $J_0(p)$ est, par construction, isogène à un quotient de $\tilde{J}(p)$ qui a un nombre fini de points rationnels (proposition I.3.2), donc $T.J_0(p)(\mathbb{Q})$ est fini, et $h_T^+(P) \in C$ le groupe cuspidal de $J_0(p)$. D'après la proposition I.3.1 (b), il suffit alors de montrer que la réduction modulo p de $h_T^+(P)$ appartient à la composante neutre de $J_0(p)_{\overline{\mathbb{F}_p}}$. Soit $\mathfrak{P}$ un idéal premier de $\mathcal{O}_K$ au-dessus de p et $R = \mathcal{O}_{K,\mathfrak{P}}$. Comme $p > 2[K:\mathbb{Q}]+1$, la proposition I.5.1 (c) nous dit que les réductions modulo $\mathfrak{P}$ de (E, A_p) et $(E/B_p, E[p]/B_p)$ appartiennent à la même composante dans $(X_0(p)_R)_{\overline{\mathbb{F}_p}}$. Alors, la structure du groupe des composantes du modèle de Néron sur R de $J_0(p)$, donnée par la proposition I.3.8, permet de voir que le diviseur $\mathrm{cl}(([E, A_p]) - ([E/B_p, E[p]/B_p]))$ est nul dans ce groupe. Il en est de même pour la différence des deux autres termes, donc la spécialisation de $h_T^+(P)$ appartient à la composante neutre de $J_0(p)_R$, d'où $h_T^+(P) = 0$. □

On peut maintenant démontrer la bonne réduction potentielle dans le cas « normalisateur de Cartan déployé ».

Théorème I.4 (Bonne réduction, cas des $\mathbb{Q}$-courbes)**.**
Soient K un corps quadratique et E une $\mathbb{Q}$-courbe de degré d définie sur K. Pour $p > 23$ premier, si $\mathbb{P}\overline{\rho}_{E,p}$ est à valeurs dans le normalisateur d'un sous-groupe de Cartan déployé de $\mathrm{PGL}(E[p])$*, alors E a potentiellement bonne réduction en tout idéal de $\mathcal{O}_K$.*

Démonstration. Soit P le point de $X_0^{\mathrm{s}}(d;p)(K)$ associé à E (proposition I.1.17) et $T \in \mathbb{T}$ tel que $T(1+w_p) = 0$, annulant $\gamma_{\mathcal{I}}$ et congru à 1 modulo $\mathcal{I}$ (lemme I.3.3). D'après la proposition I.5.4, on a $h_T^+(P) = 0$ et on peut alors appliquer la méthode de Mazur.

Commençons par la caractéristique p : soit $\mathfrak{P}$ un idéal de $\mathcal{O}_K$ au-dessus de p, montrons que E a potentiellement bonne réduction en $\mathfrak{P}$. Supposons que E a réduction potentiellement multiplicative modulo $\mathfrak{P}$. Comme $p > 23$, d'après la proposition I.5.1 (a) et (b), la réduction modulo $\mathfrak{P}$ de P appartient à la partie lisse de $X_0^{\mathrm{s}}(d;p)$. Comme toutes les pointes de cette partie lisse sont au-dessus de la pointe ∞_p dans $X_{\mathrm{split}}(p)_{\mathbb{F}_p}$ par l'étude de la fibre en p de $X_{\mathrm{split}}(p)$ et que d est sans facteur carré, on peut appliquer une involution d'Atkin-Lehner étendue de $X_0(d)$ à $X_0^{\mathrm{s}}(d;p)$ pour supposer que la réduction modulo $\mathfrak{P}$ de P est exactement $(\infty,\infty)_p$. Mais alors, en prenant $T \in \mathbb{T}\backslash p\mathbb{T}$ annulant $\gamma_{\mathcal{I}}$, h_T^+ est une immersion formelle en $(\infty,\infty)_p$ qui envoie P sur 0 d'après la proposition I.5.4. Par la proposition I.2.13, on a donc $P = (\infty,\infty)$, ce qui est absurde.

Supposons maintenant que λ est un idéal premier de $\mathcal{O}_K$ au-dessus du nombre premier $\ell \neq p$ et que E a réduction potentiellement multiplicative modulo λ. Nous allons tout d'abord étudier les réductions des pointes non rationnelles pour prouver que P ne peut se réduire en ces pointes modulo ℓ. L'image d'une pointe de $X_0^{\mathrm{sC}}(d;p)$ par h est 0 si et seulement si cette pointe est au-dessus de ∞ ou 0 dans $X_{\mathrm{sp.Car.}}(p)$, et $2[0]-2[\infty]$ sinon. En effet, si c est une pointe de $X_0^{\mathrm{sC}}(d;p)$ au-dessus d'une pointe différente de ∞ ou 0 dans $X_{\mathrm{sp.Car.}}(p)$(c'est-à-dire une pointe non rationnelle), son image par $\pi_1 : X_0^{\mathrm{sC}}(d;p) \to X_0(p)$ (définie par $\pi_1((E, C_d, A_p, B_p)) = (E, A_p)$) est la pointe 0 de

$X_0(p)$ (cela se vérifie sur les morphismes de surfaces de Riemann). En conséquence, par définition de h, $h(c) = 2[0] - 2[\infty]$.

Le groupe cyclique C de $J_0(p)(\mathbb{Q})$ engendré par $[0] - [\infty]$ est non nul pour $p > 13$ (proposition I.3.2), et d'après le lemme de spécialisation de Raynaud ([Maz78], Proposition 1.1) est étale hors de la caractéristique 2 et en caractéristique 2, a éventuellement un seul élément non trivial dont la spécialisation est nulle, qui est le point de 2-torsion de C. Ainsi, si on avait $h(c) = 0$ modulo ℓ, on aurait, si $\ell > 2$, $2 = 0$ modulo $\mathrm{num}((p-1)/12)$ d'où $p = 3$ ou 7, et si $\ell = 2$, $4 = 0$ modulo $\mathrm{num}((p-1)/12)$, d'où $p = 3, 5, 7$ ou 13, ce qu'on a exclu.

Prenons maintenant $T \in \mathbb{T}\backslash\ell\mathbb{T}$ annulant $\gamma_{\mathcal{I}}$ et inversible modulo $\mathcal{I}$ (lemme I.3.3). Alors, T est injective sur la réduction de C sur ℓ donc $h_T^+(c)$ est non nul modulo ℓ pour toute pointe c de $X_0^{\mathrm{s}}(d;p)$ non au-dessus de ∞ par le travail précédent. Or, $h_T^+(P)$ est nul modulo ℓ car il est nul dans $J_0(p)(\mathbb{Q})$ par la proposition I.5.4. Cela impose que P se réduit en une pointe au-dessus de ∞ dans $X_{\mathrm{split}}(p)$ modulo λ. En appliquant une involution d'Atkin-Lehner, on peut donc supposer que P se réduit en $(\infty,\infty)_\ell$ modulo λ. Comme h_T^+ est une immersion formelle en ce point, on en déduit par la proposition I.2.13 que $P = (\infty,\infty)$, ce qui est absurde, donc E a potentiellement bonne réduction en tout idéal premier de $\mathcal{O}_K$. □

Remarque I.5.3. Nous pouvons maintenant analyser en quoi cette preuve diffère du cas Borel :
- Il n'existe pas de morphisme canonique de la courbe modulaire $X_{\mathrm{split}}(p)$ dans $J_0(p)$. Pour les définitions, on doit en revenir à $X_{\mathrm{sp.Car.}}(p)$ puis passer au quotient.
- Le groupe d'Atkin-Lehner ne suffit pas pour se ramener en une seule pointe, mais le raisonnement permet de se ramener à (∞,∞). Cela rend la preuve plus lisible, mais n'est pas nécessaire : on peut prouver qu'en toutes les autres pointes, le morphisme est bien une immersion formelle (voir dans cette idée la proposition 2.5 de [Mom84]).
- Le fait que l'image par h_T du point concerné soit nulle nous évite d'avoir à utiliser la proposition I.2.15 (et contourne donc les difficultés posées habituellement en caractéristique 2 et 3), mais on utilise à répétition le lemme de spécialisation de Raynaud dans les résultats préliminaires.

I.6 Le cas des normalisateurs de sous-groupes de Cartan non déployés

En plus des méthodes utilisées dans les cas précédents, nous allons ici devoir employer des nouveaux outils, de nature analytique. Nous allons comprendre comment se ramener dans cette section à des estimations analytiques, qui seront faites dans le chapitre II. Commençons par des rappels et des notations sur la courbe modulaire associée au cas non déployé et sa jacobienne. Notons J la jacobienne de la courbe tordue $X = X_0^{\mathrm{ns}}(d;p)^K$ décrite dans la section I.1.4.

Définition I.6.1 (Algèbre de Hecke)**.** Pour tout entier n premier à dp, on peut définir pour $X = X^{\mathrm{ns}}(d;p)$ ou $X = X_0(dp^2)$ l'opérateur de Hecke T_n grâce aux morphismes

$$\pi_1, \pi_2 : X_0(n) \times_{X(1)} X \to X,$$

où π_1 est le morphisme d'oubli de la n-structure et $\pi_2 = \pi_1 \circ w_n$ où w_n est l'involution d'Atkin-Lehner de degré n sur $X_0(n) \times_{X(1)} X$ relevant celle de degré n sur $X_0(n)$.

Ces morphismes définissent une correspondance sur X et donc un endomorphisme noté $T_n = {\pi_2}_*\pi_1^*$ de $\mathrm{Jac}(X)$ défini sur $\mathbb{Q}$. De plus, leur action sur les q-développements des formes modulaires est la même que celle décrite dans le paragraphe I.2.2. Les T_n, avec n premier à dp, engendrent donc une algèbre d'opérateurs sur $\mathrm{Jac}(X)$ qui est commutative et un $\mathbb{Z}$-module libre de type fini (comme dans le cas $X = X_0(p)$).

On note alors $\mathbb{T}'$ l'algèbre engendrée par ces opérateurs, à laquelle on adjoint le groupe W des involutions d'Atkin-Lehner de degré divisant d sur X.

On peut ramener l'étude de la jacobienne de $X_0^{\mathrm{ns}}(d;p)$, en apparence mystérieuse, à une jacobienne beaucoup plus familière, grâce au lemme suivant, énoncé dans [Ell04].

Lemme I.6.2. *Pour tout p premier et tout d premier à p sans facteur carré, il existe une isogénie*

$$\alpha : \mathrm{Jac}(X^{ns}(d;p)) \to J_0'(dp^2)/((w_{p^2}-1)J_0'(dp^2)) \oplus J_0(d)$$

équivariante pour l'action de $\mathbb{T}'$, avec $J_0'(dp^2)$ la partie p-nouvelle de $J_0(dp^2)$.

Démonstration. Pour $d=1$, c'est un résultat dû à [Che00] et au moins implicite dans [DM97], mais nous allons pour cette généralisation utiliser un résultat de De Smit et Edixhoven. On note $G=\mathrm{GL}_2(\mathbb{F}_p)$, B un sous-groupe de Borel, C un sous-groupe de Cartan déployé de normalisateur N et C' un sous-groupe de Cartan non déployé de normalisateur N' de G .

$$\begin{aligned} M^C \oplus M^G \oplus M^G &\cong M^{C'} \oplus M^B \oplus M^B \\ M^N \oplus M^G &\cong M^{N'} \oplus M^B. \end{aligned}$$

Pour appliquer ce résultat, on fixe $\mathcal{C}=\mathbb{Q}\otimes\mathcal{A}$ avec $\mathcal{A}$ la catégorie des variétés abéliennes sur $\mathbb{Q}$ munies d'une action de $\mathbb{T}'$, c'est-à-dire telle que ses objets sont de telles variétés abéliennes, et $\mathrm{Hom}_{\mathcal{C}}(A,B)=\mathbb{Q}\otimes\mathrm{Hom}_{\mathcal{A}}(A,B)$ pour tous $A,B\in\mathcal{C}$. En particulier, deux objets de $\mathcal{C}$ sont isomorphes si et seulement s'ils sont isogènes dans $\mathcal{A}$. On choisit alors $M=\mathrm{Jac}(X_0(d)\times_{X(1)}X(p))$, de sorte que $M^H=\mathrm{Jac}(X_0(d)\times_{X(1)}X(p)/H)$ pour tout sous-groupe H de G. En effet, pour tout objet A de $\mathcal{C}$, il est clair que $\mathrm{Hom}_{\mathcal{C}}(A,M)^H=\mathrm{Hom}(A,\mathrm{Jac}(X_0(d)\times_{X(1)}X(p))^H)$ pour l'action naturelle de G sur $\mathrm{Jac}(X)$ provenant de celle sur $X(p)$, et alors $\mathrm{Jac}(X_0(d)\times_{X(1)}X(p))^H$ est une sous-variété abélienne de $\mathrm{Jac}(X_0(d)\times_{X(1)}X(p))$ dont l'espace cotangent en 0 est exactement $S_2(\Gamma_0(d)\cap\Gamma(p))^H$, c'est-à-dire $S_2(\Gamma_0(d)\cap\Gamma_H)$ avec Γ_H l'image réciproque de H par la projection de $\mathrm{SL}_2(\mathbb{Z})$ dans $\mathrm{SL}_2(\mathbb{F}_p)$. Comme $\mathrm{Jac}(X_0(d)\times_{X(1)}X(p))^H$ contient clairement la jacobienne associée à ce sous-groupe de congruences, on a l'égalité $\mathrm{Jac}(X_0(d)\times_{X(1)}X(p))^H=\mathrm{Jac}(X_0(d)\times_{X(1)}X(p)/H)$.

En conséquence, on a $M^G=J_0(d)$ car $X(p)/G=X(1)$, $M^B=J_0(dp)$ car $X(p)/B=X_0(p)$ par définition, $M^C=\mathrm{Jac}(X_0^{\mathrm{sC}}(d;p))$, $M^N=\mathrm{Jac}(X_0^{\mathrm{s}}(d;p))$ et $M^{N'}=\mathrm{Jac}(X^{\mathrm{ns}}(d;p))$.

Or, $X_0(p^2)$ et $X^{\mathrm{sp.Car.}}(p)$ sont isomorphes. En effet, à tout couple (E,C_{p^2}) avec E une courbe elliptique et C_{p^2} un sous-groupe cyclique d'ordre p^2 de E, on peut associer $E'=E/C_{p^2}[p]$. On note $A_p=E[p]/C_{p^2}[p]\subset E'$ qui est cyclique de cardinal p, et $B_p=C_{p^2}/C_{p^2}[p]$ qui l'est également. Alors, $A_p\neq B_p$ et on obtient donc un triplet $(E',A_p,B_p)\in X^{\mathrm{sp.Car.}}(p)$. Réciproquement, pour un triplet $(E',A_p,B_p)\in X^{\mathrm{sp.Car.}}(p)$, on considère $E=E'/A_p$ et C_{p^2} le groupe cyclique de cardinal p^2 de E qui est le noyau de l'isogénie naturelle $E\to E'\to E/B_p$ de degré p^2. On vérifie directement que ces applications sont réciproques, et que l'isomorphisme $\varphi : X_0(p^2)\to X^{\mathrm{sp.Car.}}(p)$ obtenu vérifie $\varphi\circ w_p=w\circ\varphi$ avec les notations de la section I.5. On en déduit que $M^C=\mathrm{Jac}(X_0(dp^2))$ et $M^N=\mathrm{Jac}(X_0(dp^2)/w_p)$, et le théorème 2 de [dSE00] nous donne donc que $J_0(dp)\oplus\mathrm{Jac}(X_0^{\mathrm{ns}}(d;p))$ et $\mathrm{Jac}(X_0(dp^2)/w_p)\oplus J_0(d)$ sont isogènes, ce qui prouve le lemme.

□

Un des points-clés de l'argument d'Ellenberg est la proposition suivante (page 8 de [Ell04]), dont nous reprenons ici la preuve en détail (avec quelques modifications).

Proposition I.6.3. *Soient $d>1$ sans facteur carré et p un nombre premier ne divisant pas d. Soit f une forme modulaire qui est soit une forme nouvelle propre de $S_2(\Gamma_0(d'p^2))$ avec d' un diviseur propre de d telle que $f_{|w_p}=f$, soit une forme nouvelle propre de $S_2(\Gamma_0(dp^2))$ telle que $f_{|w_p}=f$ et $f_{|w_d}=-f$.*

Alors, il existe un morphisme surjectif $\pi_f : J_0(dp^2)_{\mathbb{Q}}\longrightarrow(A_f)_{\mathbb{Q}}$ où A_f est la variété quotient de $J_0(d'p^2)$ (resp. $J_0(dp^2)$) associée à f, tel que $\pi_f\circ w_d=-\pi_f$ et $\pi_f\circ w_p=\pi_f$.

Démonstration. Le deuxième cas de définition de f est évident par construction de la variété quotient, traitons donc le premier.

Commençons par des résultats et notations provenant de [AL70]. Pour tout $n\in\mathbb{N}_{>0}$ sans facteur carré, soit A_n l'opérateur qui à une forme modulaire f de poids $2k$ de niveau quelconque sur le demi-plan de Poincaré associe $f_{|A_n}$ définie par

$$f_{|A_n}(\tau)=n^k f(n\tau).$$

Pour tout $m \in \mathbb{N}$ et tout $f \in S_2(\Gamma_0(m))$, $f_{|A_n} \in S_2(\Gamma_0(mn))$ ([AL70], Lemme 2). Si m est premier à n, on a également un morphisme rationnel de variétés abéliennes $B_n : J_0(mn) \to J_0(m)$, provenant par fonctorialité d'Albanese du morphisme de dégénérescence $X_0(mn) \to X_0(m)$ qui à une classe d'isomorphisme (E, C_m, C_n) associe $(E/C_n, (C_m + C_n)/C_n)$. L'action de ce morphisme sur les espaces cotangents est donc un morphisme entre $S_2(\Gamma_0(m))$ et $S_2(\Gamma_0(mn))$, qui se trouve être exactement $\frac{1}{n}A_n$ avec A_n défini ci-dessus.

Pour tout q premier divisant n, à tout diviseur δ positif de n on associe δ_q qui vaut δq si q ne divise pas δ, et δ/q sinon. Cela définit clairement une involution entre les diviseurs de n n'ayant aucun point fixe (car n est sans facteur carré), et le lemme 26 de [AL70] nous donne en particulier la formule suivante pour toute forme modulaire f sur $S_2(\Gamma_0(m))$:

$$\left(f_{|A_\delta}\right)_{|w_q} = (f)_{|A_{\delta_q}}.$$

On en déduit par égalité sur les espaces cotangents en 0 que $\delta B_\delta \circ w_q = \delta_q B_{\delta_q}$. Alors, pour tout choix de fonction ε qui à tout diviseur δ de n associe $\varepsilon_\delta = \pm 1$, l'application

$$I_\varepsilon = \sum_{\delta|n} \delta \varepsilon_\delta B_\delta : J_0(mn) \to J_0(m)$$

vérifie, pour tout diviseur premier q de n :

$$I_\varepsilon \circ w_q = \sum_{\delta|n} \delta_q \varepsilon_\delta B_{\delta_q}.$$

Comme n est sans facteur carré, l'ensemble de ses diviseurs s'identifie naturellement à $(\mathbb{Z}/2\mathbb{Z})^r$ avec r le nombre de ses facteurs premiers, considérons donc ε comme un morphisme de groupes de $(\mathbb{Z}/2\mathbb{Z})^r$ dans $\{\pm 1\}$ via cette identification, et donc entièrement déterminé par ses valeurs en chaque diviseur premier q de n. Alors, pour chacun de ceux-ci, on a

$$I_\varepsilon \circ w_q = \varepsilon_q I_\varepsilon$$

de sorte que $I_\varepsilon \circ w_d$ est égal à $\pm I_\varepsilon$ suivant la parité du nombre de diviseurs premiers de δ tels que $\varepsilon_q = -1$.

Pour une forme f satisfaisant la deuxième hypothèse de l'énoncé, on fixe $m = d'p^2$ et $n = d/d'$ dans les notations précédentes. La projection naturelle $\pi'_f : J_0(d'p^2) \to A_f$ vérifie $\pi'_f \circ w_{d'} = \pm\pi'_f$ et $\pi'_f \circ w_p = \pi'_f$, on adapte alors le choix de ε pour que $\pi_f := \pi'_f \circ I_\varepsilon$ satisfasse $\pi_f \circ w_d = -\pi_f$ et $\pi_f \circ w_p = \pi_f$ (c'est possible car d est premier à p). De plus, l'application I_ε est surjective car l'application cotangente associée (de $S_2(\Gamma_0(d'p^2))$ vers $S_2(\Gamma_0(dp^2))$) est une combinaison linéaire des opérateurs A_δ, δ divisant n, or sur le q-développement, on a, si $g(\tau) = \sum_{k=1}^{+\infty} a_k(g) e^{2i\pi k\tau}$,

$$g_{|A_m}(\tau) = m \sum_{k=1}^{+\infty} a_k(g) e^{2i\pi km\tau}.$$

En particulier, si g est non nul, en prenant le premier coefficient $a_k(g)$ non nul, on observe que pour le plus petit m divisant d/d' dont le coefficient est non nul dans la combinaison linéaire, le mk-ième coefficient de l'image de g par l'application cotangente est exactement $a_k(g)$ qui est donc non nul. Ainsi, I_ε est bien surjective car son application cotangente est injective, et l'application π_f l'est donc également.

□

Prenons maintenant f qui vérifie les hypothèses de la proposition I.6.3 et à coefficients entiers pour chacun des développements en les pointes. Grâce au morphisme α du lemme I.6.2, on en déduit un morphisme surjectif rationnel $\pi_f : J_0^{\mathrm{ns}}(d;p)_{\mathbb{Q}} \to A_{f\mathbb{Q}}$ tel que $\pi_f \circ w_d = -\pi_f$ (la condition $\pi_f \circ w_p = \pi_f$ servant à passer au quotient par $w_p - 1$).

Soit maintenant K un corps quadratique. On note χ son caractère quadratique, σ son automorphisme non trivial et $A_f \otimes \chi$ la tordue de A_f pour $K/\mathbb{Q}$ relativement à -1 (c'est-à-dire la variété

abélienne définie sur $\mathbb{Q}$, et isomorphe à A_f par un isomorphisme j défini sur K tel que ${}^\sigma j = -j$). De même, on note J la tordue de $\mathrm{Jac}(X_0^{\mathrm{ns}}(d;p))$ pour $K/\mathbb{Q}$ relativement à l'automorphisme w_d. On a alors un diagramme commutatif de variétés abéliennes

$$\begin{array}{ccc} J_0^{ns}(d;p) & \xrightarrow{\pi'_f} & A'_f \\ \downarrow{\scriptstyle i} & & \downarrow{\scriptstyle j} \\ J & \xrightarrow{\psi'_f} & A'_f \otimes \chi \end{array}$$

dont toutes les flèches sont des morphismes définis sur K, avec i un isomorphisme de variétés abéliennes défini sur K tel que ${}^\sigma i = i \circ w_d$, et ψ'_f défini pour faire commuter ce diagramme. Il est défini sur K et

$${}^\sigma\psi'_f = {}^\sigma(j \circ \pi'_f \circ i^{-1}) = -j \circ \pi'_f \circ w_d \circ i^{-1} = -j \circ (-\pi'_f) \circ i^{-1} = \psi'_f.$$

Ainsi, ψ'_f est défini sur $\mathbb{Q}$.

Rappelons que la courbe $X_{\mathrm{nonsplit}}(p)$ est une courbe modulaire sur $\mathbb{Q}$, admettant exactement $(p-1)/2$ pointes à valeurs dans $\mathbb{Q}(\zeta_p+\zeta_p^{-1})$ et toutes conjuguées par l'action de $\mathrm{Gal}(\mathbb{Q}(\zeta_p+\zeta_p^{-1})/\mathbb{Q})$ ([Ser97], Annexe A.5).

On note ∞ la pointe infinie usuelle de $X_0(d)$, (∞,∞) la pointe infinie naturelle de $X_0^{\mathrm{ns}}(d;p)$, R_0 l'anneau des entiers de $\mathbb{Q}(\zeta_p+\zeta_p^{-1})$ et $R = R_0[\frac{1}{2p}]$. La courbe $X_0^{\mathrm{ns}}(d;p)$ a un modèle propre sur R, noté $X_0^{\mathrm{ns}}(d;p)_R$, qui est lisse sur $R[\frac{1}{d}]$ (c'est prouvé dans la section 5 de [DM97] dans des cas particuliers, mais également vrai par exactement les mêmes arguments en général). De plus, chaque pointe de $X_0^{\mathrm{ns}}(d;p)$ est définie sur R_0 et pour chaque idéal premier λ au-dessus d'un diviseur de d, la pointe $(\infty,\infty)_\lambda$ appartient à la partie lisse de ce modèle. On définit le morphisme d'Albanese usuel $h : X_0^{\mathrm{ns}}(d;p)_{\mathbb{Q}} \to J_0^{\mathrm{ns}}(d;p)_{\mathbb{Q}}$ qui envoie (∞,∞) sur 0, et par abus de notation, on note à nouveau

$$h : X_0^{\mathrm{ns}}(d;p)_R^{\mathrm{lisse}} \to J_0^{\mathrm{ns}}(d;p)_R$$

son extension par propriété de Néron au modèle de Néron de $J_0^{\mathrm{ns}}(d;p)$ sur R. Dans ce cas, on n'aura besoin que d'une immersion formelle, forcément introduite par des méthodes légèrement différentes des deux premiers cas.

Proposition I.6.4 (Immersion formelle)**.** *Avec les notations précédentes, le morphisme $\pi'_f \circ h$ est une immersion formelle le long de la section $(\infty,\infty)_R$.*

Démonstration. C'est le lemme 3.8 de [Ell04], qui est basé sur le lemme 8.2 de [DM97]. □

Théorème I.5 (Bonne réduction)**.**

Soit K un corps quadratique imaginaire et $d > 1$ un entier sans facteur carré.

Supposons que pour un certain $p \geq 29$ premier, il existe f vérifiant les hypothèses de la proposition I.6.3 telle que $A_f \otimes \chi$ est de rang zéro sur $\mathbb{Q}$. Alors, pour toute $\mathbb{Q}$-courbe E de degré d définie sur K telle que $\mathbb{P}\overline{\rho}_{E,p}$ est à image dans un normalisateur d'un sous-groupe de Cartan non déployé de $\mathrm{PGL}(E[p])$, E a potentiellement bonne réduction en tout idéal premier de $\mathcal{O}_K$.

Remarque I.6.1. Ici, pas d'amélioration qualitative ni quantitative : ce résultat est implicite dans la preuve de la proposition 3.9 de [Ell04]. Nous allons néanmoins en réécrire la preuve ci-dessous. Il est à noter que le même résultat serait valable pour un corps quadratique réel (remarque 3.7 de [Ell04]), mais il faut que pour χ le caractère quadratique associé à K, on ait $\chi(d) = 1$ pour que l'hypothèse sur f soit vérifiable, et le premier cas de la proposition I.6.3 est impossible pour des raisons de signe de l'équation fonctionnelle (section II.1). Nous ne le mentionnons qu'en remarque ici car la méthode de Runge de la section I.7 est de toute façon inapplicable (la *condition de Runge* pour $X_0(q)$ avec q premier ne pouvant pas être utilisée pour les corps quadratiques réels, théorème III.6), et on ne peut alors pas avoir de résultat d'uniformité pour les $\mathbb{Q}$-courbes quadratiques réelles. Cependant, la section II.4 fera les calculs dans le cas quadratique réel sous l'hypothèse $\chi(d) = 1$ et d premier, à la fois parce qu'ils offrent une variante intéressante, et qu'ils permettent de retrouver analytiquement un résultat de quotient de rang zéro de [DM97].

Le lemme suivant traite spécialement le cas de la bonne réduction au-dessus de p.

Lemme I.6.5 (Supersingularité automatique en p)**.**
Soit K un corps de nombres, p un nombre premier et E une courbe elliptique définie sur K telle que $\rho_{E,p} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(E[p])$ est à valeurs dans le normalisateur d'un sous-groupe de Cartan non déployé de $\mathrm{GL}(E[p])$. Alors, si $p > 12[K:\mathbb{Q}]+1$, E a potentiellement bonne réduction (supersingulière) en tout idéal premier de $\mathcal{O}_K$ au-dessus de p.

Démonstration. Supposons que E a potentiellement mauvaise réduction ou bonne réduction ordinaire en un idéal premier $\mathfrak{p}$ au-dessus de p dans $\mathcal{O}_K$. Notons $K_\mathfrak{p}$ le complété du corps K pour la valuation associée à $\mathfrak{p}$. Il existe une extension K' de degré divisant 4 ou 6 de $K_\mathfrak{p}$ telle que E a réduction semi-stable sur $K'_\mathfrak{P}$. D'après les propositions I.1.12 et I.1.13, l'image du groupe d'inertie I' de K' est de cardinal multiple de $(p-1)/\mathrm{pgcd}(e', p-1)$ avec $e' = v(p)$ l'indice de ramification absolu de K', et constituée d'éléments de la forme

$$\begin{pmatrix} * & * \\ 0 & 1 \end{pmatrix}.$$

En particulier, tous ces éléments admettent la valeur propre 1. Par hypothèse, c'est le cas pour au plus deux éléments de l'image de $\rho_{E,p}$, car celle-ci est dans le normalisateur d'un sous-groupe de Cartan non déployé. On en déduit donc que $(p-1)/e' \leq 2$ et par construction, $e' \leq 6[K:\mathbb{Q}]$ d'où $p \leq 12[K:\mathbb{Q}]+1$, ce qui prouve le lemme par contraposée. □

Remarque I.6.2. Le cas « normalisateur de Cartan déployé » excluait la supersingularité (proposition I.5.1 (a)), c'est pour ainsi dire le contraire qui se produit dans le cas « normalisateur de Cartan non déployé ». C'est aussi pourquoi nous n'aurions pu utiliser une immersion formelle pour nous occuper de la caractéristique p, les réductions étant alors en des points non lisses, sur lesquels h n'est pas définie.

Nous pouvons maintenant prouver le théorème I.5.

Démonstration. Soient p un nombre premier et E et une $\mathbb{Q}$-courbe de degré d vérifiant les hypothèses du théorème.

Les pointes de $X_0^{\mathrm{ns}}(d;p)$ au-dessus de ∞ dans $X_0(d)$ ont pour corps de définition $\mathbb{Q}(\zeta_p+\zeta_p^{-1})$ et $\mathrm{Gal}(\mathbb{Q}(\zeta_p+\zeta_p^{-1})/\mathbb{Q})$ agit transitivement sur l'ensemble de ces pointes. Comme K est imaginaire, son intersection avec ce corps est $\mathbb{Q}$. En conséquence, les pointes de $X_0^{\mathrm{ns}}(d;p)$ au-dessus de ∞ dans $X_0(d)$ sont toutes conjuguées par $\mathrm{Gal}(\overline{K}/K)$. Prenons ℓ un nombre premier différent de p (le lemme précédent s'étant chargé des idéaux premiers au-dessus de p) et λ un idéal premier de $K(\zeta_p+\zeta_p^{-1})$ au-dessus de ℓ. Supposons par l'absurde que E a potentiellement mauvaise réduction modulo λ, c'est-à-dire que P se réduit en une pointe modulo λ. Alors, les involutions d'Atkin-Lehner de degré divisant d agissant transitivement sur les pointes de $X_0(d)$, on peut supposer que P se réduit en une pointe au-dessus de ∞ dans $X_0(d)$ (cela ne change pas la nature de la réduction semi-stable). Par la transitivité de l'action de $\mathrm{Gal}(\overline{K}/K)$, on peut en fait choisir λ tel que P se réduit en (∞,∞) modulo λ. On remarque alors que pour qu'un K-point de $X_0^{\mathrm{ns}}(d;p)$ se réduise en (∞,∞) modulo $\lambda\cap\mathcal{O}_K$, le corps résiduel $\mathcal{O}_K/\lambda\cap\mathcal{O}_K$ doit contenir $\zeta_p+\zeta_p^{-1}$ avec ζ_p une racine p-ième primitive de l'unité dans $\overline{\mathbb{F}_\ell}$. Ce corps résiduel étant $\mathbb{F}_\ell$ ou $\mathbb{F}_{\ell^2}$, cela implique que $\zeta_p+\zeta_p^{-1} \in \mathbb{F}_{\ell^2}$, mais alors $\zeta_p \in \mathbb{F}_{\ell^4}$, donc $p|\ell^4-1$. C'est impossible pour $\ell = 2,3$ lorsque $p \geq 29$, ce qui écarte d'ores et déja les cas $\ell = 2,3$, et on suppose dorénavant que $\ell \geq 5$. Prenant f qui vérifie les hypothèses de la proposition I.6.3, l'application

$$\psi'_f \circ h : X_0^{\mathrm{ns}}(d;p)^K_{|R} \longrightarrow A_f \otimes \chi_{|R}$$

est une immersion formelle le long de la section $(\infty,\infty)_R$ par la proposition I.6.4 (la conjugaison par des isomorphismes n'y change rien). Par ailleurs, le point $g(P)$ est un point de torsion dans $A_f \otimes \chi(K(\zeta_p+\zeta_p^{-1}))$: en effet, par le théorème de Drinfeld-Manin, le groupe des diviseurs cuspidaux

de $J_0^{\mathrm{ns}}(d;p)$ est de n-torsion pour un certain n. Alors, pour tout $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, et tout point Q de $X_0^{\mathrm{ns}}(d;p)^K$,

$$\begin{aligned} {}^\tau(\psi'_f \circ h)(Q) - (\psi'_f \circ h)(Q) &= \psi'_f([{}^\tau Q] - [Q] + [(\infty,\infty)] - [{}^\tau(\infty,\infty)]) \\ &= \psi'_f([{}^\tau Q] - [Q]) + \psi'_f([(\infty,\infty)] - [{}^\tau(\infty,\infty)]) \end{aligned}$$

car ψ'_f est défini sur $\mathbb{Q}$, donc

$$n\,{}^\tau(\psi'_f \circ h)(Q) = n(\psi'_f \circ h)(Q)$$

pour tout Q, c'est-à-dire que $n \cdot \psi'_f \circ h$ est défini sur $\mathbb{Q}$. En particulier, il envoie les points $\mathbb{Q}$-rationnels sur les points $\mathbb{Q}$-rationnels. Par construction, P est un point $\mathbb{Q}$-rationnel de $X_0^{\mathrm{ns}}(d;p)^K$ et $A_f \otimes \chi$ est de rang zéro sur $\mathbb{Q}$, donc $n \cdot (\psi'_f \circ h(P))$ est bien de torsion, donc $\psi'_f \circ h(P)$ également. Ensuite, l'indice de ramification absolu de R_λ est au plus 2 car λ n'est pas au-dessus de p (et $\mathbb{Q}(\zeta_p + \zeta_p^{-1})$ n'est ramifiée qu'en p), et il est donc strictement inférieur à $\ell - 1$ car $\ell > 3$. On peut donc appliquer la proposition I.2.16 qui nous dit que $P = (\infty,\infty)$, ce qui est absurde, et E a donc potentiellement bonne réduction modulo tout idéal premier de $\mathcal{O}_K$. □

Il reste maintenant à trouver dans quelles conditions on peut trouver de telles variétés abéliennes tordues de rang zéro sur $\mathbb{Q}$.

C'est l'objet du théorème II.1, auquel est consacré le chapitre II, dont on écrit la conséquence ici.

Théorème I.6. *Soit K un corps quadratique imaginaire de discriminant $-D$ et de caractère χ, et $d > 1$ un entier sans facteur carré. Alors, pour tout nombre premier $p > 50D^{1/4}\log D$ ne divisant pas D, il existe une forme modulaire f vérifiant les hypothèses du théorème I.5. Ainsi, d'après le théorème I.5, pour toute $\mathbb{Q}$-courbe E de degré d définie sur K telle que $\mathbb{P}\overline{\rho}_{E,p}$ est à image dans un normalisateur d'un sous-groupe de Cartan non déployé de $\mathrm{PGL}(E[p])$, E a potentiellement bonne réduction en tout idéal premier de $\mathcal{O}_K$.*

I.7 La méthode de Runge pour $X_0(p)$ avec p premier

Dans cette section (qui reprend la partie 3 de [LF]), nous allons donner une borne sur la hauteur des points entiers de $X_0(p)$ avec p un nombre premier fixé (attention, on utilisera cette borne pour un diviseur premier du degré de la $\mathbb{Q}$-courbe, et non pas pour p le nombre premier de la non-surjectivité de la représentation galoisienne).

La méthode de Runge sera décrite en toute généralité dans le chapitre III, aussi bien pour des courbes que pour des variétés de dimension supérieure. Cependant, nous donnons ici son application aux courbes $X_0(p)$ car elle est particulièrement simple, et quasiment autocontenue.

Tout d'abord, $X_0(p)$ n'a que deux pointes notées c_∞ et c_0, qui sont respectivement l'orbite de ∞ et de 0 par $\Gamma_0(p)$. Elles sont définies sur $\mathbb{Q}$ ainsi que $X_0(p)$. La méthode de Runge ne peut s'appliquer qu'à un ensemble de places de cardinal strictement plus petit que le nombre d'orbites de pointes (section III.3 pour les courbes modulaires, section III.2 pour le cas général), et donc uniquement à K un corps quadratique imaginaire et S constitué de l'unique place infinie de K. On note

$$X_0(p)(\mathcal{O}_K) := \{P \in X_0(p)(K) \,|\, j(P) \in \mathcal{O}_K\}$$

(les points de cet ensemble sont dit $\mathcal{O}_K$-entiers). Nous allons exhiber par des moyens plus directs que la section III.3 les unités modulaires sur $X_0(p)$.

Définition I.7.1. Soit p un nombre premier fixé. Avec la notation $q_\tau = e^{2i\pi\tau}$, on définit la fonction holomorphe g sur $\mathcal{H}$ par

$$g(\tau) = \frac{\Delta(\tau)}{\Delta(p.\tau)} = q_\tau^{1-p} \prod_{\substack{n=1 \\ (p,n)=1}}^{+\infty} (1-q_\tau^n)^{24}.$$

avec Δ la forme modulaire discriminant sur $\mathcal{H}$. En tant que quotient de deux formes modulaires de poids 12 pour $\Gamma_0(p)$, elle définit une fonction (modulaire) sur $X_0(p)$ qu'on note U.

La fonction U a les propriétés suivantes.

Proposition I.7.2. (*a*) *Pour tout $\tau \in \mathcal{H}$:*

$$g(-1/\tau) = p^{12} g^{-1}(\tau/p) = p^{12} q^{(p-1)/p} \prod_{n=1}^{\infty} \left(\frac{1-q^n}{1-q^{n/p}} \right)^{24}.$$

(*b*) *Pour w_p l'involution d'Atkin-Lehner sur $X_0(p)$, on a $U \circ w_p = p^{12} U^{-1}$.*
(*c*) *Le diviseur de U sur $X_0(p)$ est porté par les pointes, plus précisément*

$$\mathrm{div}(U) = (p-1)([c_0] - [c_\infty]).$$

(*d*) *Les fonctions U et $p^{12}U^{-1}$ sont $\mathbb{Q}$-rationnelle sur $X_0(p)$ et entières sur $\mathbb{Z}[j]$.*

Démonstration. Le (*a*) implique (*b*), car pour tout $\tau \in \mathcal{H}$ d'image P dans $X_0(p)$, on a $g(\tau) = U(P)$ et $-1/(p\tau)$ a pour image $w_p(P)$ dans $X_0(p)$ par définition. Il suffit pour prouver le (*a*) d'écrire que pour tout $\tau \in \mathcal{H}$, par définition de g,

$$g(-1/\tau) = \frac{\Delta(-1/\tau)}{\Delta(-p/\tau)} = \frac{\tau^{12}\Delta(\tau)}{p^{-12}\tau^{12}\Delta(\tau/p)} = p^{12} g^{-1}(\tau/p),$$

car Δ est modulaire de poids 12 pour $\mathrm{SL}_2(\mathbb{Z})$. Comme celle-ci ne s'annule pas sur $\mathcal{H}$, le diviseur de U est bien porté par c_∞ et c_0, et le q-développement de g en $i\infty$ montre que l'ordre de U en c_∞ est $-(p-1)$ (celui-ci est non ramifié sur $X(1)$), et l'ordre en c_0 est nécessairement son opposé, d'où le (*c*). Ensuite, les fonctions U et U^{-1} sont $\mathbb{Q}$-rationnelle sur $X_0(p)$ comme quotients de deux formes modulaires $\mathbb{Q}$-rationnelles sur $X_0(p)$. Enfin, prouvons l'intégralité. Rappelons que

$$\mathrm{SL}_2(\mathbb{Z}) = \Gamma_0(p) \cup \bigcup_{k \in \mathbb{Z}} \Gamma_0(p) \cdot \begin{pmatrix} 0 & 1 \\ -1 & -k \end{pmatrix}. \tag{I.6}$$

En effet, pour tout $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, soit $\gamma \cdot \infty = \gamma' \cdot \infty$ with $\gamma' \in \Gamma_0(p)$, et alors $\gamma \in \Gamma_0(p)$ car le stabilisateur de ∞ dans $\mathrm{SL}_2(\mathbb{Z})$ est inclus dans $\Gamma_0(p)$, soit $\gamma \cdot \infty = \gamma' \cdot 0$ avec $\gamma' \in \Gamma_0(p)$. La matrice $w = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ envoie ∞ sur 0, donc

$$(\gamma' w)^{-1} \gamma = \pm \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

pour un entier k, car $(\gamma' w)^{-1}\gamma \cdot \infty = \infty$, ce qui prouve (I.6). On en déduit que pour tout $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, le q-développement de $g_{|\gamma}$ est une série formelle en $q_\tau^{1/p}$ à coefficients entiers algébriques, car si $\gamma \in \Gamma_0(p) \cdot \begin{pmatrix} 0 & 1 \\ -1 & -k \end{pmatrix}$,

$$g_{|\gamma}(\tau) = g(-1/(\tau+k)) = p^{12} e^{2i\pi(p-1)k/p} {q_\tau}^{(p-1)/p} \prod_{n=1}^{\infty} \left(\frac{1-q_\tau^n}{1-e^{2i\pi nk/p} q_\tau^{n/p}} \right)^{24}$$

par le (*a*), donc ce q_τ-développement est à coefficients dans $\mathbb{Z}[e^{2i\pi/p}] \subset \overline{\mathbb{Z}}$. On en déduit par la proposition III.3.15 que U est entier sur $\mathbb{Z}[j]$, et de même pour $p^{12}U^{-1}$ (où le p^{12} est pour compenser l'inverse du développement ci-dessus, en ∞ le développement de U^{-1} est à coefficients entiers). □

Remarque I.7.1. Avec les notations précédentes pour les unités modulaires, on voit tout de suite que

$$g = \prod_{a=1}^{p-1} g_{(\frac{a}{p},0)}^{12p},$$

ce qui aurait permis de retrouver toute la proposition précédente sauf le (b) et le (d). À noter que la proposition III.3.22 aurait seulement obtenu que $p^{12p}U^{-1}$ est entière sur $\mathbb{Z}[j]$, ce qui est moins fort que le (d).

Lemme I.7.3. *Pour tout $P \in Y_0(p)(\mathcal{O}_K)$, $U(P)$ est un élément non nul de $\mathcal{O}_K$ tel que*

$$0 \leq \log|U(P)| \leq 12\log p.$$

Démonstration. Comme U est $\mathbb{Q}$-rationnelle et entière sur $\mathbb{Z}[j]$ (proposition I.7.2 (d)), $U(P) \in \mathcal{O}_K$ et est non nul car U ne s'annule pas sur $Y_0(p)$. De même, $p^{12}U(P)^{-1}$ l'est, et comme K est un anneau quadratique imaginaire, pour tout $\alpha \in \mathcal{O}_K$ non nul, $\log|\alpha| \geq 0$, donc

$$0 \leq \log|U(P)| \leq 12\log p.$$

□

On définit w sur $\mathcal{H}$ par $w(\tau) = -1/\tau$ et g_0 sur $\mathcal{H}$ par $g_0 := g \circ w$.

Le lemme suivant, dit « lemme de localisation près des pointes », permettra de réduire nos calculs pour Runge à des estimations des deux fonctions g et g_0 près des pointes.

Lemme I.7.4. *On note $\mathcal{D}$ le domaine fondamental « usuel » de $\mathcal{H}$ pour $\mathrm{SL}_2(\mathbb{Z})$, c'est-à-dire le domaine délimité par le triangle géodésique de $\mathcal{H} \cup \infty$ pour la topologie hyperbolique, de sommets $\infty, e^{i\pi/3}, e^{2i\pi/3}$.*

Pour tout $P \in Y_0(p)(\mathbb{C})$, il existe $\tau \in \mathcal{D} + \mathbb{Z}$ tel que τ ou $-1/\tau$ est au-dessus de P via la projection canonique $\mathcal{H} \to Y_0(p)(\mathbb{C})$. Dans le premier cas, on dit que P est près de c_∞, et alors $j(P) = j(\tau)$ et $U(P) = g(\tau)$. Dans le second cas, on dit que P est près de c_0, et alors $j(P) = j(\tau)$ et $U(P) = g_0(\tau)$.

Démonstration. Soit $P \in Y_0(p)(\mathbb{C})$. On choisit $\tau_0 \in \mathcal{H}$ au-dessus de P, il existe donc $\beta \in \mathrm{SL}_2(\mathbb{Z})$ tel que $\beta \cdot \tau_0 = \tau_1 \in \mathcal{D}$. Ce τ_1 n'est plus nécessairement au-dessus de P à moins que $\beta \in \Gamma_0(p)$ (et on choisit alors $\tau = \tau_1$ pour le lemme, et P est près de c_∞). Si $\beta \notin \Gamma_0(p)$, grâce à (I.6), on peut écrire

$$\beta^{-1} = \gamma \cdot w \cdot \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

pour un $k \in \mathbb{Z}$ et $\gamma \in \Gamma_0(p)$. Ainsi, $\tau = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \cdot \tau_1 \in \mathcal{D} + \mathbb{Z}$, et $w.\tau = \gamma^{-1} \cdot \tau_0$ est au-dessus de P. On dit alors que P est près de c_0. □

On aura besoin du lemme suivant pour estimer assez précisément les q-développements de g et g_0.

Lemme I.7.5. *Pour tout $r \in]0,1[$ et tout $q \in \mathbb{C}$ tel que $|q| \leq r$,*

$$\sum_{n=1}^{+\infty} |\log|1-q^n|| \leq \frac{-\log(1-r)}{r(1-r)}|q|.$$

Pour tout $q \in \mathbb{C}$ tel que $|q| < 1$,

$$\sum_{n=1}^{+\infty} |\log|1-q^n|| \leq \frac{\pi^2}{6\log|q^{-1}|}. \tag{I.7}$$

Démonstration. La première inégalité est une conséquence directe de l'inégalité triangulaire et du principe du maximum. La seconde est basée sur la preuve du lemme 3.4 de [BP11a] (mais à un terme borné près, qu'on souhaite enlever). Il suffit de la montrer pour $q = x$ un réel entre 0 et 1 par comparaison avec les modules. Si $x \geq 1/2$, cela provient de la première inégalité avec $r = 1/2$ (c'est-à-dire que celle-ci est plus fine que (I.7) dans ce cas). On suppose donc que $x \geq 1/2$. Pour tout $\tau \in \mathcal{H}$, avec $\Delta = \eta^{24}$, on a

$$\Delta(-1/\tau) = \tau^{12}\Delta(\tau).$$

Étant donné $x \geq 1/2$, on fixe $\tau_x = \log(x)/(2i\pi) \in \mathcal{H}$ de sorte que $q_\tau = x$ et

$$Q_x = \exp(2i\pi/\tau) \in]0, 1[.$$

Le développement en ∞ de Δ nous donne alors

$$\log|Q_x| + 24\sum_{n=1}^{+\infty}\log(1-Q_x^n) = 12\log|\tau| + \log x + 24\sum_{n=1}^{+\infty}\log(1-x^n).$$

Ceci se traduit en

$$-\sum_{n=1}^{+\infty}\log(1-x^n) = -\frac{\pi^2}{6\log x} + \frac{\log x}{24} + \frac{\log(|\tau|)}{2} - \sum_{n=1}^{\infty}\log(1-Q_x^n).$$

Il reste donc à démontrer que la somme des trois derniers termes de droite est négative. Le second est négatif car $x < 1$ et le troisième terme est majoré par -1 car $x \geq 1/2$, donc

$$\log|\tau| = \log(\log(x)/(2\pi)) \leq \log(\log(2)/(2\pi)) \leq -1.$$

Le dernier terme est majoré par 10^{-23} car $Q_x \leq \exp(-4\pi^2/\log 2) \leq 10^{-24}$, et on utilise la première inégalité avec $r = 10^{-24}$. □

Proposition I.7.6. *Pour tout $\tau \in \mathcal{D} + \mathbb{Z}$,*

$$\begin{aligned} |\log|g(\tau)| + (p-1)\log|q_\tau|| &\leq 25|q_\tau|. \\ |\log|g_0(\tau)| - \frac{p-1}{p}\log|q_\tau|| &\leq \frac{4\pi^2 p}{\log|q_\tau^{-1}|} + 12\log(p). \end{aligned}$$

Démonstration. D'après le q-développement de g et le lemme I.7.5 appliqué à $r = 0.005$,

$$|\log|g_\infty(\tau)| + (p-1)\log|q_\tau|| = 24\left|\sum_{\substack{n\geq 1\\(p,n)=1}}\log|1-q_\tau^n|\right| \leq -24\frac{\log(0.995)}{0.995\cdot 0.005}|q_\tau| \leq 25|q_\tau|$$

car $|q_\tau| \leq 0.005$ lorsque $\tau \in \mathcal{D} + \mathbb{Z}$. Pour g_0, on utilise l'autre inégalité du lemme I.7.5, et on obtient

$$|\log|g_0(\tau)| - \frac{p-1}{p}\log|q_\tau| - 12\log(p)| = 24\left|\sum_{\substack{n\geq 1\\(p,n)=1}}\log|1-q_\tau^{n/p}|\right| \leq \frac{4\pi^2}{\log|q_\tau^{-1/p}|} = \frac{4\pi^2 p}{\log|q_\tau^{-1}|}.$$

□

Nous pouvons maintenant donner notre borne sur le j-invariant.

Proposition I.7.7. *Soit K un corps quadratique imaginaire. Pour tout nombre premier p et tout $P \in Y_0(p)(\mathcal{O}_K)$,*

$$\log|j(P)| < 2\pi\sqrt{p} + 6\log(p) + 8.$$

Démonstration. Soit $\tau \in \mathcal{D} + \mathbb{Z}$ associé à P par le lemme de localisation près des pointes. Si $\log|j(P)| < 2\pi\sqrt{p}$, il n'y a rien à prouver. Sinon, $|j(\tau)| > 3500$ donc $\log|j(\tau)| \leq \log|q_\tau^{-1}| + \log(2)$ par la proposition III.3.19 (d). Il faut maintenant borner $|q_\tau^{-1}|$. Si P est près de c_∞, d'après le lemme I.7.3 on a $\log|g(\tau)| = \log|U(P)| \leq 12\log p$. D'après la proposition I.7.6, on obtient

$$\log|q_\tau^{-1}| \leq \frac{25|q_\tau| + 12\log p}{p-1} \leq 2\pi\sqrt{p} \tag{I.8}$$

après un petit calcul utilisant que $|q_\tau| \leq 0.005$.

Si P est près de c_0, $\log|g_0(\tau)| = \log|U(P)| \geq 0$ par le lemme I.7.3. Par la proposition I.7.6, on obtient cette fois que

$$\left(\frac{p-1}{p}\right)\log|q_\tau^{-1}| \leq \frac{4\pi^2 p}{\log|q_\tau^{-1}|} + 12\log(p),$$

donc

$$\log|q_\tau^{-1}| \leq \frac{2\pi p}{\sqrt{p-1}} + \frac{6p\log(p)}{(p-1)} \leq 2\pi\sqrt{p} + 6\log(p) + 7 \tag{I.9}$$

car $p \geq 2$, après un petit calcul sur les termes du reste. Dans les deux cas, (I.8) ou (I.9) nous donnent la majoration annoncée. □

Remarque I.7.2. On a ici affaire à une fonction du niveau beaucoup plus petite que dans le cas général (où elle est au mieux linéaire, théorème III.6), au point que le terme quadratique vient du terme d'erreur de l'estimation de $g_0(\tau)$, et non pas du contrôle de $U(P)$.

I.8 Théorèmes d'isogénie et résultat principal

Le moment est venu de conclure la preuve du théorème I.1 grâce à un théorème d'isogénie basé sur les résultats de [GR14], et énoncé dans la section 5 de [LF], que nous reproduisons quasiment à l'identique ici. Il se formule avec la hauteur stable de Faltings, notée $h_\mathcal{F}$, que nous ne redéfinirons pas (voir la section 3.3 de [GR14]), mais liée à la hauteur du j-invariant par l'inégalité suivante : pour toute courbe elliptique E définie sur un corps de nombres K, on a

$$h_\mathcal{F}(E) \leq \frac{1}{12}h(j_E) - 1.28, \tag{I.10}$$

où h est la hauteur logarithmique absolue de $j(E) \in K$, c'est-à-dire

$$h(j(E)) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K}[K_v:\mathbb{Q}_v]\max(0,\log(|j(E)|_v))$$

avec M_K l'ensemble des places de K. Cette inégalité est le lemme 7.9 de [GR14], en prenant en compte le fait que pour les notations de ce lemme, on a $h_\mathcal{F}(E) = h(E) - (1/2)\log\pi$.

Le résultat suivant est une conséquence directe de la section 7.3 de [GR14], mais non formulée telle quelle dans cet article, c'est pourquoi nous en donnons une preuve succincte ci-dessous.

Proposition I.8.1. *Soit K un corps de nombres.*

Soit E une courbe elliptique sans multiplication complexe et B une surface abélienne toutes les deux sur K, avec $\psi : B \to E \times E$ une isogénie sur K. Supposons (hypothèse $(*)$) que pour tout plongement $\sigma : K \to \mathbb{C}$, si $\Omega_{E,\sigma}$ et $\Omega_{B,\sigma}$ sont les réseaux de périodes de E_σ et B_σ, le sous-réseau $\mathrm{d}\psi(\Omega_{B,\sigma})$ de $\Omega_{E,\sigma}^2$ contient un élément (ω_1,ω_2) de $\Omega_{E,\sigma}$ qui est une $\mathbb{Z}$-base de $\Omega_{E,\sigma}$. Alors,

$$\deg(\psi) \leq 10^7[K:\mathbb{Q}]^2(\max\{h_\mathcal{F}(E), 985\} + 4\log[K:\mathbb{Q}])^2.$$

Démonstration. La borne donné ici est exactement celle du théorème 1.4 de [GR14], car le calcul est presque exactement le même que celui de la preuve de ce théorème, nous allons expliquer pourquoi. Pour tout plongement $\sigma : K \to \mathbb{C}$, il existe une norme canonique $\|\cdot\|_\sigma$ (provenant d'une

polarisation principale de E) sur l'espace tangent $t_{E,\sigma}$, qui contient $\Omega_{E,\sigma}$. On fixe un plongement σ_0 tel qu'il existe une base (ω_1, ω_2) de Ω_{E,σ_0} qui est minimale parmi toutes les bases minimales pour tous les réseaux de périodes possibles $\Omega_{E,\sigma}$, c'est-à-dire que

$$\|\omega_1\|_{\sigma_0} = \max_{\sigma} \min_{\substack{\omega \in \Omega_{E,\sigma} \\ \omega \neq 0}} \|\omega\|_{\sigma},$$

et $\omega_2 = \tau\omega_1$ pour τ dans le domaine fondamental de Siegel de $\mathcal{H}$ (de sorte que $y = \mathrm{im}(\tau)$ est minimal parmi tous les choix de plongements et de bases minimales pour chacun des plongements). Ce choix de σ_0 est donc le même que dans la section 7.3 de [GR14]. Nous allons maintenant confondre toutes les variétés abéliennes sur K considérées comme leurs extensions des scalaires à $\mathbb{C}$ via ce σ_0, et dorénavant omettre tout mention des plongements dans nos notations. On peut composer ψ par un isomorphisme de $E \times E$ tel que $\mathrm{d}\psi(\Omega_E)$ contient la base (ω_1, ω_2) choisie précédemment, car on a supposé que ψ vérifie l'hypothèse $(*)$. Soit alors $A = E \times E \times B$ et une période $\omega = (\omega_1, \omega_2, \chi)$ de $\Omega_E^2 \times \Omega_B$, où $\chi \in \Omega_B$ est tel que $\mathrm{d}\psi(\chi) = (\omega_1, \omega_2)$. La sous-variété abélienne minimale A_ω de A contenant ω dans son espace tangent en 0 est alors

$$A_\omega = \{(\psi(z), z), z \in B\}.$$

En effet, l'inclusion $A_\omega \subset \{(\psi(z), z), z \in B\}$ est évidente, et réciproquement, la projection canonique de A_ω dans $E \times E$ donne une sous-variété abélienne de $E \times E$ contenant (ω_1, ω_2) dans son réseau des périodes. Or, E est une courbe elliptique sans multiplication complexe, donc l'anneau des endomorphismes de $E \times E$ est $M_2(\mathbb{Z})$, et aucune sous-variété abélienne stricte de $E \times E$ ne peut contenir (ω_1, ω_2) dans son espace tangent. Ceci prouve que la dimension de A_ω est au moins 2, d'où l'égalité ci-dessus.

Maintenant, la variété abélienne A_ω est canoniquement isomorphe à B, et la projection de A_ω vers $E \times E$ est une isogénie de degré Δ. On a obtenu toutes les données nécessaires pour reprendre les calculs de la section 7.3 de [GR14] et obtenir la borne de la proposition, à deux légères différences près. La première est que le plongement σ_0 peut être réel ou complexe, et la borne peut changer si σ_0 est réel (la section 7.3 suppose qu'il est complexe). Pour éviter ce problème, on peut considérer $K' = K(i)$ et reprendre les calculs du départ avec K' au lieu de K : cela concorde avec la preuve de la section 7.3, qui nécessitait aussi une extension de degré 2 éventuelle de K. Ensuite, le calcul des pentes dans le lemme 7.6 de [GR14] est légèrement différent, mais il suffit d'utiliser que B est isogène à $E \times E$ avec isogénie de degré Δ pour donner exactement la même inégalité. □

La proposition I.8.1 va nous permettre de prouver une version explicite du théorème de surjectivité de Serre, énoncée ci-dessous. Rappelons que le vocabulaire et les résultats principaux sur les sous-groupes particuliers de $\mathrm{GL}_2(\mathbb{F}_p)$ (qu'on utilisera sans mention spécifique) sont donnés dans la section I.1.2.

Théorème I.7. *Soit K un corps de nombres, et E une courbe elliptique sur K sans multiplication complexe. Soit $\mathcal{B}$ (resp. $\mathcal{C}$) un ensemble de nombres premiers p tels que l'image de $\rho_{E,p}$ est incluse dans un sous-groupe de Borel (resp. le normalisateur d'un groupe de Cartan, qu'il soit déployé ou non) de $\mathrm{GL}(E[p])$. Alors, on a l'inégalité suivante :*

$$\prod_{p \in \mathcal{B}} p \prod_{q \in \mathcal{C}} \frac{q^2}{4} \leq 10^7 [K : \mathbb{Q}]^2 \left(\max\{h_{\mathcal{F}}(E), 985\} + 4\log[K : \mathbb{Q}] + 4|\mathcal{C}|\log(2)\right)^2.$$

En particulier, la représentation $\rho_{E,p}$ est surjective pour tout nombre premier

$$p > 10^7 [K : \mathbb{Q}]^2 \left(\max\{h_{\mathcal{F}}(E), 985\} + 4\log[K : \mathbb{Q}]\right)^2$$

non ramifié dans K.

Tout d'abord, nous allons prouver un lemme technique permettant plus tard d'assurer que l'isogénie ψ qu'on construira vérifie bien l'hypothèse $(*)$ de la proposition I.8.1.

Lemme I.8.2. *Soit p un nombre premier, V un $\mathbb{F}_p$-espace vectoriel de dimension 2, et $\underline{v}$ une base de V. Alors, pour tout $g \in \mathrm{GL}(V)$, la forme quadratique*

$$Q_{\underline{v}} : x \longmapsto \det_{\underline{v}}(x, g.x)$$

est surjective dans $\mathbb{F}_p$ si g est semi-simple, à moins que ce ne soit une homothétie (et alors elle est nulle).

Démonstration. Tout d'abord, si g est simple, la forme $Q_{\underline{v}}$ ne peut avoir aucun vecteur isotrope, et est donc surjective, (en regardant sa forme dans une base orthogonale). Sinon, rappelons que pour tous $x, y \in V$ et toutes bases $\underline{v}$ et $\underline{v'}$,

$$\det_{\underline{v'}}(x, y) = \det_{\underline{v'}}(\underline{v}) \cdot \det_{\underline{v}}(x, y)$$

donc deux formes quadratiques $Q_{\underline{v}}$ provenant de deux choix de bases différents sont proportionnelles. En conséquence, $Q_{\underline{v}}$ est surjective si et seulement si $Q_{\underline{v'}}$ l'est pour une autre base de notre choix. Si g est semi-simple mais pas simple ni une homothétie, il existe une base $\underline{v} = (v_1, v_2)$ dans laquelle g est diagonale, à valeurs propres distinctes λ et μ. Alors, pour $x = x_1 v_1 + x_2 v_2$, on a

$$Q_{\underline{v}}(x) = (\mu - \lambda) x_1 x_2,$$

donc $Q_{\underline{v}}$ est bien surjective. □

Passons maintenant à la preuve du théorème I.7.

Démonstration. Nous allons prouver la première inégalité, et expliquer ensuite pourquoi elle implique la seconde. Soit K' une extension de K de degré $2^{|\mathcal{C}|}$, choisie de telle sorte que pour tout $q \in \mathcal{C}$, l'image de $\rho_{E,q}$ restreinte à $\mathrm{Gal}(\overline{K}/K')$ est incluse dans un sous-groupe de Cartan (déployé ou non), et pas seulement dans un normalisateur. Soit $n_{\mathcal{B}} = \prod_{p \in \mathcal{B}} p$ et $n_{\mathcal{C}} = \prod_{q \in \mathcal{C}} q$. Alors, l'inégalité à prouver se réécrit comme

$$n_{\mathcal{B}} n_{\mathcal{C}}^2 \leq 10^7 [K' : \mathbb{Q}]^2 \left((\max\{h_{\mathcal{F}}(E), 985\} + 4 \log[K' : \mathbb{Q}]) \right)^2.$$

Il suffit donc de trouver une surface abélienne B et une isogénie $\psi : B \to E \times E$ sur K' telles que ψ est de degré $n_{\mathcal{B}} n_{\mathcal{C}}^2$ et satisfait l'hypothèse $(*)$ pour appliquer la proposition I.8.1. Nous allons maintenant construire ψ et B.

Pour tout $p \in \mathcal{B}$, on choisit une $\mathbb{F}_p$-droite C_p de $E[p]$ fixée par $\rho_{E,p}$, et on note $G_p = C_p \times E[p]$, qui est un sous-groupe de $E[p]^2$ de cardinal p^3. Pour tout $q \in \mathcal{C}$, on choisit un élément g_q du sous-groupe de Cartan associé à q qui n'est pas une homothétie (un tel élément est toujours semi-simple), et on pose $G_q = \{(x, g_q \cdot x), x \in E[q]\} \subset E[q]^2$. Ce group est de cardinal q^2 et stable par l'action diagonale du sous-groupe de Cartan (car ceux-ci sont commutatifs), donc défini sur K' dans $E \times E$. On considère alors

$$G = \bigoplus_{p \in \mathcal{B}} G_p \oplus \bigoplus_{q \in \mathcal{C}} G_q \subset E[n_{\mathcal{B}} n_{\mathcal{C}}]^2.$$

C'est un groupe de cardinal $n_{\mathcal{B}}^3 n_{\mathcal{C}}^2$, défini sur K', et on pose B la variété abélienne quotient $B = (E \times E)/G$ et $\varphi : E \times E \to B$ la projection canonique de noyau G. Comme G est un sous-groupe de $E[n_{\mathcal{B}} n_{\mathcal{C}}]^2$, il existe une isogénie $\psi : B \to E \times E$ telle que

$$\psi \circ \varphi = [n_{\mathcal{B}} n_{\mathcal{C}}],$$

c'est-à-dire la multiplication par $n_{\mathcal{B}} n_{\mathcal{C}}$ sur $E \times E$. Cette isogénie ψ est définie sur K', et par multiplicativité des degrés on obtient $\deg(\psi) = n_{\mathcal{B}} n_{\mathcal{C}}^2$. Il reste donc seulement à prouver que ψ vérifie l'hypothèse $(*)$.

Soit $\sigma : K' \to \mathbb{C}$ un plongement. Pour ce plongement, on identifie E_σ avec le quotient de son espace tangent $t_{E,\sigma}$ par son réseau des périodes $\Omega_{E,\sigma}$ (on sous-entendra la dépendance en σ

systématiquement dans la suite). Si π est la projection $t_E \times t_E \to E \times E$, le réseau $\Omega = \pi^{-1}(G)$ de t_E^2 définit une variété abélienne quotient $t_E \times t_E/\Omega$ isomorphe à B. Avec ces plongements, on a le diagramme commutatif

$$\begin{array}{ccccc} t_E \times t_E & \xrightarrow{\mathrm{Id}} & t_E \times t_E & \xrightarrow{n_{\mathcal{B}} n_{\mathcal{C}}} & t_E \times t_E \\ \downarrow{\scriptstyle \pi} & & \downarrow & & \downarrow{\scriptstyle \pi} \\ E \times E & \xrightarrow{\varphi} & B & \xrightarrow{\psi} & E \times E. \end{array}$$

Voyons maintenant pourquoi $\Omega' = n_{\mathcal{B}} n_{\mathcal{C}} \Omega \subset \Omega_E \times \Omega_E$ contient une base de Ω_E. Soit (e_1, e_2) une base de Ω_E.

On considère l'image de Ω' dans $(\Omega_E/p\Omega_E)^2$ pour $p \in \mathcal{B} \cup \mathcal{C}$. Après multiplication par $1/p$, cette image s'identifie avec un sous-groupe de $((\Omega_E/p)/(\Omega_E))^2$, c'est-à-dire $E[p]^2$. D'après les définitions de Ω et Ω', l'image de Ω'/p dans $E[p]^2$ est $((n_{\mathcal{B}} n_{\mathcal{C}})/p)G_p = G_p$ (la partie première à p de Ω est envoyé dans Ω_E^2 lorsqu'elle est multipliée par $n_{\mathcal{B}} n_{\mathcal{C}}/p$). Une base de $E[p]^2$ est $\pi(e_1/p, e_2/p)$ et on identifie par la suite $E[p]$ à ${\mathbb{F}_p}^2$ pour ce choix de base.

Pour $p \in \mathcal{B}$, soit un vecteur non nul (a, b) de C_p et $(c, d) \in {\mathbb{F}_p}^2$ tel que $ad - bc = 1 \mod p$. Alors, le vecteur $((a, b), (c, d))$ appartient à G_p et est de déterminant 1 pour notre choix de base. Pour $q \in \mathcal{C}$, d'après le lemme I.8.2, il existe $x \in E[q]$ tel que $\det_{\pi(e_1/q, e_2, q)}(x, g_q \cdot x) = 1$, et le sous-groupe G_q de $E[q]^2$ contient donc un élément de déterminant 1 pour notre choix de base.

On vient donc de prouver que pour tout $p \in \mathcal{B} \cup \mathcal{C}$, il existe une matrice $\gamma_p = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{F}_p)$ telle que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} \overline{e_1} \\ \overline{e_2} \end{pmatrix} \in \Omega'/(p\Omega_E)^2$$

Le morphisme de spécialisation $\mathrm{SL}_2(\mathbb{Z}) \to \prod_{p \in \mathcal{B} \cup \mathcal{C}} \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ étant surjectif ([Lan02], Chapitre XIII, Exercise 18) , il existe donc $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ tel que

$$\gamma \cdot \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \in \Omega' + (n_{\mathcal{B}} n_{\mathcal{C}})\Omega_E^2 \subset \Omega',$$

car Ω contient Ω_E^2. Un tel élément $\gamma \cdot {}^t(e_1, e_2)$ est par construction une base de Ω_E, donc Ω' contient bien une base de Ω_E, ce qui conclut la preuve de la première inégalité.

Soit maintenant p un nombre premier non ramifié dans K et tel que $\rho_{E,p}$ n'est pas surjective. L'image de $\rho_{E,p}$ est alors incluse dans un sous-groupe de Borel, le normalisateur d'un sous-groupe de Cartan, ou un sous-groupe exceptionnel. Dans le cas « Borel », on obtient

$$p \leq 10^7 [K : \mathbb{Q}]^2 \left(\max\{h_{\mathcal{F}}(E), 985\} + 4 \log[K : \mathbb{Q}]\right)^2.$$

Dans le cas « normalisateur de Cartan » (déployé ou non) on obtient

$$p \leq 2 \cdot 10^{3.5} [K : \mathbb{Q}] \left(\max\{h_{\mathcal{F}}(E), 985\} + 4 \log[K : \mathbb{Q}] + 4 \log(2)\right).$$

Dans le cas exceptionnel, on a $p \leq 30[K : \mathbb{Q}] + 1$ par la proposition I.1.15. Le maximum des trois bornes est donc obtenu pour le cas Borel, et nous donne la seconde inégalité. □

Remarque I.8.1. Certains articles, par exemple [Coj05], donnent d'autres versions (non totalement explicites) du théorème de surjectivité de Serre, utilisant le conducteur N_E de E au lieu de la hauteur de Faltings stable. On peut cependant remarquer que, d'après le théorème 3 de [Coj05] (supposant la « conjecture du degré »), le théorème I.7 implique que pour E définie sur $\mathbb{Q}$, la représentation $\rho_{E,p}$ est surjective pour $p \gg \log(N_E)^2$.

On donne maintenant l'application du théorème I.7 aux $\mathbb{Q}$-courbes.

Corollaire I.8.1. *Soit E une $\mathbb{Q}$-courbe sans multiplication complexe, de degré sans facteur carré d), et sur un corps quadratique K. Alors :*

- Si $\mathbb{P}\overline{\rho}_{E,p}$ est réductible pour un certain nombre premier p ne divisant d,

$$dp \leq 10^7[K:\mathbb{Q}]^2(\max\{h_{\mathcal{F}}(E), 985\} + 4\log[K:\mathbb{Q}])^2.$$

- Si $\mathbb{P}\overline{\rho}_{E,p}$ a son image incluse dans le normalisateur d'un sous-groupe de Cartan (déployé ou non) pour un certain nombre premier p ne divisant pas d,

$$dp^2 \leq 4 \cdot 10^7[K:\mathbb{Q}]^2(\max\{h_{\mathcal{F}}(E), 985\} + 4\log(2[K:\mathbb{Q}]))^2.$$

On peut maintenant énoncer le théorème I.1 plus en détail.

Théorème I.8. *Soit K un corps quadratique imaginaire de discriminant $-D_K$. Pour toute $\mathbb{Q}$-courbe stricte E définie sur K, de degré $d(E)$, et tout nombre premier p ne divisant pas $d(E)$,*

- *Si $p \geq 2 \cdot 10^{13}$, la représentation $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas à valeurs dans un sous-groupe de Borel.*
- *Si $p \geq 10^7$, la représentation $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas à valeurs dans le normalisateur d'un sous-groupe de Cartan déployé.*
- *Si $p \geq \max(10^7, 50D_K^{1/4}\log(D_K))$ et ne divise pas D_K, la représentation $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas à valeurs dans le normalisateur d'un sous-groupe de Cartan non déployé.*
- *Si $p \geq 67$, la représentation $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas incluse dans un sous-groupe exceptionnel.*

Démonstration. Tout d'abord, rappelons qu'on peut supposer que $d(E)$ est sans facteur carré (théorème I.2). Le cas exceptionnel vient de la proposition I.1.15. Pour les autres cas, on utilise les théorèmes I.3 et I.6 et la proposition I.4. On en déduit qu'avec les bornes données, si $\mathbb{P}\overline{\rho}_{E,p}$ n'est pas surjective, $j(E) \in \mathcal{O}_K$. Soit maintenant d_0 le plus petit diviseur premier de d. D'après la proposition I.7.7, on a alors

$$\log|j(E)| \leq 2\pi\sqrt{d_0} + 6\log(d_0) + 8 \leq 2\pi\sqrt{d} + 6\log(d) + 8.$$

Comme $j(E) \in \mathcal{O}_K$ et K est quadratique imaginaire, on a $h_{\mathcal{F}}(E) \leq (\log|j(E)|)/12 - 1$ d'après l'inégalité (I.10). En combinant cette borne supérieure sur $\log|j(E)|$ avec les bornes inférieures données par le corollaire I.8.1, on obtient alors chacun des cas du théorème. □

Concluons ce chapitre avec quelques remarques sur les étapes de la preuve. Tout d'abord, il est probable qu'on puisse utiliser la méthode de Mazur pour toute $\mathbb{Q}$-courbe centrale sur un corps de nombres donné dans les cas « Borel » et « normalisateur de Cartan déployé », avec des bornes dépendant seulement du degré de ce corps. En effet, les preuves dans ces deux cas s'appuient essentiellement sur le fait que $X_0(d) \to X(1)$ est non ramifié en la pointe ∞ et ramifié en toutes les autres pointes. On pourrait donc utiliser les involutions d'Atkin-Lehner pour obtenir une immersion formelle vérifiant les bonnes conditions, ainsi qu'un quotient de rang zéro sur $\mathbb{Q}$. De plus, le travail technique sur le groupe des composantes est fait en toute généralité sur le degré du corps dans la proposition I.3.8. Pour ce qui est du cas « normalisateur de Cartan non déployé », il est possible que l'astuce d'Ellenberg pour fabriquer l'immersion formelle puisse être appliqué à certains corps de degré plus grand.

Au sujet de la méthode de Runge, remarquons qu'une $\mathbb{Q}$-courbe centrale de degré d (où d a r facteurs premiers distincts) est définie sur un corps de degré 2^r, et on peut appliquer la méthode de Runge à $X_0(d)(\mathcal{O}_K)$ à moins que K ne soit totalement réel, avec sans doute un exposant $1/2$ en $\log|j(E)|$ comme dans le cas quadratique. Cependant, on connaît beaucoup moins d'exemples de $\mathbb{Q}$-courbes sur des corps de degré plus grand, il n'est donc pas sûr que le problème d'uniformité sur ces corps soit pertinent (exemple I.1.2).

II

Estimation de moyennes de fonctions L

« Tout le monde peut être intelligent »
– José Mourinho

II.1 Préparation de l'estimation

Nous devons maintenant trouver des variétés abéliennes quotients de $\mathrm{Jac}(X_0^{\mathrm{ns}}(d;p)))$ dont la tordue est de rang zéro sur $\mathbb{Q}$, d'après la section I.6. Les résultats de ce chapitre sont pour une grande partie tirés de l'Annexe de [LF] et de [LF15]. Les notations de cette section sont les suivantes :

K est un corps quadratique et D la valeur absolue de son discriminant.

χ est le caractère quadratique associé, de conducteur D.

$\mathcal{H}$ est le demi-plan de Poincaré.

Pour $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R})$ et $z \in \mathcal{H}$, on note

$$\gamma \cdot z := \frac{az+b}{cz+d}, \qquad j_\gamma(z) := cz + d.$$

On en déduit immédiatement les formules suivantes :

$$\mathrm{Im}(\gamma \cdot z) = \frac{(\det \gamma)\,\mathrm{Im}\, z}{|j_\gamma(z)|^2}, \qquad \gamma \cdot z = \frac{a}{c} - \frac{\det \gamma}{c j_\gamma(z)}, \qquad j_{\gamma\gamma'}(z) = j_\gamma(\gamma' \cdot z) j_{\gamma'}(z). \tag{II.1}$$

Elles seront parfois utilisées dans cette partie sans mention spécifique. Pour toute fonction holomorphe f sur $\mathcal{H}$ et toute matrice $\gamma \in \mathrm{GL}_2^+(\mathbb{R})$, on définit la fonction $f_{|\gamma}$ sur $\mathcal{H}$ par

$$f_{|\gamma}(z) := \frac{\det \gamma}{j_\gamma(z)^2} f(\gamma \cdot z),$$

et ceci définit une action à droite de $\mathrm{GL}_2^+(\mathbb{R})$ sur ces fonctions, telle que f est une fonction modulaire de poids 2 pour $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ si et seulement si f est Γ-invariante.

Pour $N \in \mathbb{N}_{>0}$, on note $S_2(\Gamma_0(N))^+$ (resp. $S_2(\Gamma_0(N))^-$, $S_2(\Gamma_0(N))^{\mathrm{old}}$, $S_2(\Gamma_0(N))^{\mathrm{new}}$) le sous-espace vectoriel des formes modulaires f de $S_2(\Gamma_0(N))$ telles que $f \circ w_N = f$ (resp. telles que $f \circ w_N = -f$, anciennes, et nouvelles).

Tous ces espaces sont munis du produit scalaire de Petersson noté $\langle \cdot, \cdot \rangle_N$, défini par

$$\langle f, g \rangle_N = \int_D \overline{f(x+iy)} g(x+iy) dx dy.$$

avec D un domaine fondamental quelconque de $\mathcal{H}$ pour l'action de $\Gamma_0(N)$. Pour cette convention, le produit scalaire de Petersson dépend du choix du sous-groupe de congruence ambiant, ainsi pour N' divisant N et $f, g \in S_2(\Gamma_0(N'))$, on a

$$\langle f, g\rangle_N = [\Gamma_0(N) : \Gamma_0(N')]\langle f, g\rangle_{N'}.$$

Pour tout sous-espace V de $S_2(\Gamma_0(N))$ et toutes formes linéaires A, B sur V, on note

$$(A, B)_V := \sum_{f\in\mathcal{F}_V} \overline{A(f)}B(f),$$

où $\mathcal{F}_V$ est une base de V orthonormale pour le produit scalaire de Petersson. Cette notation définit bien un produit scalaire sur V^* indépendant du choix de $\mathcal{F}_V$.

Pour tout $m \in \mathbb{N}$, on note a_m la fonctionnelle $f \mapsto a_m(f)$ le m-ième coefficient du q-développement de f, et $L_\chi : f \mapsto L(f \otimes \chi, 1)$. Par souci de concision, nous allons noter $(a_m, L_\chi)_N$ (resp. $(a_m, L_\chi)_N^+$) le produit scalaire de a_m et L_χ sur $V = S_2(\Gamma_0(N))$ (resp. $V = S_2(\Gamma_0(N))^+$) et de même pour les autres sous-espaces de $S_2(\Gamma_0(N))$.

Pour tout diviseur positif Q de N tel que $\mathrm{pgcd}(Q, N/Q) = 1$, on note W_Q toute matrice de la forme suivante :

$$W_Q := \begin{pmatrix} Qx & y \\ Nz & Qt \end{pmatrix}, \quad x, y, z, t \in \mathbb{Z}, \quad \det W_Q = Q.$$

Alors, $W_Q\Gamma_0(N) = \Gamma_0(N)W_Q$ et ne dépend pas du choix de W_Q, de sorte que pour toute forme modulaire $f \in S_2(\Gamma_0(N))$, $f_{|W_Q}$ ne dépend pas du choix de la matrice W_Q, et on appelle *involution d'Atkin-Lehner de degré Q sur $S_2(\Gamma_0(N))$* l'opérateur correspondant sur cet espace (renoté w_Q pour insister sur son aspect canonique). Pour $\varepsilon = \pm 1$, on note $S_2(\Gamma_0(N))^{\varepsilon Q}$ le sous-espace vectoriel de $S_2(\Gamma_0(N))$ constitué des formes modulaires f telles que $f_{|W_Q} = \varepsilon f$, ainsi $S_2(\Gamma_0(N))^+ = S_2(\Gamma_0(N))^{+_N}$. Pour plus de détails sur les propriétés de base de ces involutions, voir la partie 2 de [AL70]).

D'après le lemme I.6.2 et la proposition I.6.3 et grâce au célèbre résultat de Kolyvagin-Logachev [KL90] généralisé par Kato [Kat04], il suffit pour vérifier les hypothèses du théorème I.5 de prouver que pour un nombre premier p assez grand, il existe : soit une forme nouvelle propre de $S_2(\Gamma_0(dp^2))$ telle que $f_{|w_d} = -f$, $f_{|w_{p^2}} = f$ et $L_\chi(f) \neq 0$, soit une forme nouvelle propre de $S_2(\Gamma_0(d))$ telle que $f_{|w_d} = -f$ et $L_\chi(f) \neq 0$, soit une forme nouvelle propre de $S_2(\Gamma_0(p^2))$ telle que $f_{|w_{p^2}} = f$ et $L_\chi(f) \neq 0$. L'objectif de cette section est donc de démontrer que L_χ est non nulle sur l'un des trois sous-espaces $S_2(\Gamma_0(dp^2))^{+_p,-_d,new}$, $S_2(\Gamma_0(d))^{-,new}$ et $S_2(\Gamma_0(p^2))^{+,new}$ pour p assez grand, car chacun de ces $\mathbb{C}$-espaces vectoriels admet une base constituée de formes modulaires propres (à coefficients réels) ([AL70], Théorème 3). Pour cela, il suffit d'établir que pour p assez grand, si $L_\chi = 0$ sur $S_2(\Gamma_0(d))^-$,

$$(a_1, L_\chi)_{dp^2}^{+_p,-_d,new} \neq 0, \quad \text{ou} \quad (a_1, L_\chi)_{p^2}^{+,new} \neq 0.$$

Pour $g \in S_2(\Gamma_0(M))$ avec M quelconque, la fonction L de g se prolonge analytiquement en 1 et pour tout $x > 0$, on a

$$L(g, 1) = \int_x^{+\infty} g(iu)du - \int_{1/(Mx)}^{+\infty} g_{|w_M}(iu)du.$$

En particulier,

$$L(g, 1) = -L(g_{|w_M}, 1),$$

donc si $g_{|w_M} = g$, $L(g, 1) = 0$, et si $g_{|w_M} = -g$,

$$L(g, 1) = 2\int_{1/\sqrt{M}}^{+\infty} g(iu)du \tag{II.2}$$

(c'est un cas particulier de l'équation fonctionnelle de la fonction L).

Si $f \in S_2(\Gamma_0(N))$ avec N premier à D, alors $f \otimes \chi \in S_2(\Gamma_0(D^2N))$ et

$$(f \otimes \chi)_{|w_{D^2N}} = \chi(-N) f_{|w_N} \otimes \chi \tag{II.3}$$

(pour plus de détails sur ces équations fonctionnelles, voir ([Bum96], § $I.5$)).

L'équation fonctionnelle de la fonction L tordue par χ pour $f \in S_2(\Gamma_0(p^2))^+$ en 1 est donc

$$L(f \otimes \chi, 1) = -\chi(-p^2) L(f_{|w_p} \otimes \chi, 1) = -\chi(-1) L(f_{|w_p} \otimes \chi, 1).$$

En conséquence, L_χ ne peut être non nulle sur $S_2(\Gamma_0(p^2))^+$ que si $\chi(-1) = -1$ c'est-à-dire que K est quadratique imaginaire. S'il l'est, on cherchera donc à prouver que $(a_1, L_\chi)_{p^2}^{+,new} \neq 0$. Les formes modulaires telles que $f_{|w_p} = -f$ contribuent toujours pour zéro vu le signe de l'équation fonctionnelle, donc

$$(a_1, L_\chi)_{p^2}^{+,new} = (a_1, L_\chi)_{p^2}^{new} = (a_1, L_\chi)_{p^2} - (a_1, L_\chi)_{p^2}^{\text{old}} = (a_1, L_\chi)_{p^2}^{+} - (a_1, L_\chi)_{p^2}^{\text{old}}. \tag{II.4}$$

Nous allons également faire les calculs pour K quadratique réel même si nous n'appliquerons pas ceux-ci à la méthode de Mazur (remarque I.6.1). Le troisième cas de la proposition I.6.3 est alors impossible, et on se rabat sur les deux premiers cas évoqués. Par le même raisonnement, pour $f \in S_2(\Gamma_0(dp^2))^{+_p,-_d}$, on a

$$L(f \otimes \chi, 1) = -\chi(-dp^2) L(f_{|w_{dp^2}} \otimes \chi, 1) = \chi(d) L(f \otimes \chi, 1),$$

et le signe est le même pour $f \in S_2(\Gamma_0(d))^-$. On peut donc trouver des quotients de rang zéro sur $\mathbb{Q}$ dans le cas quadratique réel seulement si $\chi(d) = 1$. Une difficulté majeure pour les calculs de ces sommes pondérées est d'isoler la contribution de la partie nouvelle dans la contribution totale de notre sous-espace. Le lemme ci-dessous la résout dans les cas qui nous intéressent.

Lemme II.1.1.

Soient p premier et χ quadratique de conducteur D premier à p.

(a) Si χ est impair,

$$(a_1, L_\chi)_{p^2}^{+,new} = (a_1, L_\chi)_{p^2} - \frac{1}{p-1}(a_1, L_\chi)_p. \tag{II.5}$$

(b) Si χ est pair et $d > 1$ est un nombre premier ne divisant pas Dp tel que $\chi(d) = 1$, soit L_χ est non nulle sur $S_2(\Gamma_0(d))^-$, soit

$$(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,new} = (a_1, L_\chi)_{dp^2}^{+_{p^2},-_d} - \frac{1}{p-1}(a_1, L_\chi)_{dp}^{\chi(p)_p,-_d}. \tag{II.6}$$

Remarque II.1.1. On verra dans les preuves des théorèmes II.1 et II.2 que le premier terme de la somme tend vers 4π (resp. 2π) quand p tend vers l'infini, alors que le second tend vers 0, ce qui mènera vers une non-annulation. Ces limites à l'infini sont dues aux formules des traces de la proposition II.2.2. L'idée originale de ce type d'estimations est due à [Duk95] et le (a) raffine le lemme 3.12 de [Ell04].

Le (b) vise à récupérer les cas quadratiques réels favorables, mais la contribution de la partie ancienne (provenant de $S_2(\Gamma_0(d))$ se révèle difficile à calculer. Heureusement, on n'a pas besoin de le faire car si cette contribution est non nulle, par la preuve du (b) il existe un quotient de rang zéro sur $\mathbb{Q}$ de la tordue de $J_0(d)$ par χ donc de J par le lemme I.6.2.

Démonstration.

On reprend les notations de la preuve de la proposition I.6.3, qui sont aussi celles de la section 2 de [AL70].

(a) D'après la formule (II.4), il reste à calculer la contribution des formes anciennes pour obtenir la formule (II.5).

Soit $f \in S_2(\Gamma_0(p))$ une forme propre à coefficients réels et V_f le sous-espace de $S_2(\Gamma_0(p^2))^{\mathrm{old}}$ engendré par $f_{|A_1}$ et $f_{|A_p}$. Soit (g_1, g_2) une base orthonormale de V_f et M la matrice de (g_1, g_2) dans la base $(f_{|A_1}, f_{|A_p})$. Alors,

$$\begin{aligned}(a_1, L_\chi)_{V_f} &= \begin{pmatrix} a_1(f_{|A_1}) & a_1(f_{|A_p})\end{pmatrix} M {}^t\overline{M} \begin{pmatrix} L_\chi(f_{|A_1}) \\ L_\chi(f_{|A_p})\end{pmatrix} \\ &= \begin{pmatrix} a_1(f_{|A_1}) & a_1(f_{|A_p})\end{pmatrix} \overline{\begin{pmatrix} \langle f_{|A_1}, f_{|A_1}\rangle_{p^2} & \langle f_{|A_1}, f_{|A_p}\rangle_{p^2} \\ \langle f_{|A_p}, f_{|A_1}\rangle_{p^2} & \langle f_{|A_p}, f_{|A_p}\rangle_{p^2}\end{pmatrix}}^{-1} \begin{pmatrix} L_\chi(f_{|A_1}) \\ L_\chi(f_{|A_p})\end{pmatrix}.\end{aligned}$$

En effet, la matrice $\overline{M}^t M)^{-1} = {}^t M^{-1} \overline{M}^{-1}$ est celle du produit scalaire dans la base $(f_{|A_1}, f_{|A_p})$ comme (g_1, g_2) est orthonormale.

Remarquons que $L_\chi(f_{|A_1}) = L_\chi(f)$ et $f_{|A_p} \otimes \chi = \chi(p)(f \otimes \chi)_{|A_p}$ grâce aux q-développements, donc

$$L_\chi(f_{|A_p}) = \chi(p) \int_0^{+\infty} f_{|A_p} \otimes \chi(iu) du = p\chi(p) \int_0^{+\infty} (f \otimes \chi)(ipu) du = \chi(p) L_\chi(f). \tag{II.7}$$

Si $f \in S_2(\Gamma_0(p))^\varepsilon$ avec $\varepsilon = \pm 1$,

$$(f \otimes \chi)_{|w_{D^2 p}} = -\chi(-p)\varepsilon(f \otimes \chi) = \chi(p)\varepsilon(f \otimes \chi)$$

d'après la formule (II.3). En conséquence, $(a_1, L_\chi)_{V_f} = 0$ si $\varepsilon = -\chi(p)$, et on suppose dorénavant que $f_{|w_p} = \chi(p) f$.

Par définition du produit scalaire de Petersson, pour $f, g \in S_2(\Gamma_0(p^2))^{\chi(p)}$ on obtient que

$$\langle f_{|A_1}, g_{|A_1}\rangle_{p^2} = [\Gamma_0(p) : \Gamma_0(p^2)] \langle f, g\rangle_p = p\langle f, g\rangle_p \tag{II.8}$$

et

$$\langle f_{|A_p}, g_{|A_p}\rangle_{p^2} = p^2 \int_{\mathcal{D}} \overline{f(px + ipy)} g(px + ipy) dx dy = \int_{p\mathcal{D}} \overline{f}(x + iy) g(x + iy) dx dy \tag{II.9}$$

avec $\mathcal{D}$ un domaine fondamental pour $\Gamma_0(p^2)$. Comme $p\mathcal{D}$ est alors un domaine fondamental pour Γ le sous-groupe des matrices de $\mathrm{SL}_2(\mathbb{Z})$ diagonales modulo p, d'indice p dans $\Gamma_0(p)$, on a $\langle f_{|A_p}, g_{|A_p}\rangle_{p^2} = p\langle f, g\rangle_p$.

Il reste à calculer $\langle f_{|A_1}, g_{|A_p}\rangle_{p^2}$, nous employons ici une méthode différente de [Ell04], remplaçant l'usage d'une fonction L de Rankin-Selberg par des résultats de [AL70]. On a

$$\begin{aligned}\langle f_{|A_1}, g_{|A_p}\rangle_{p^2} &= p \int_{\mathcal{D}} \overline{f(x + iy)} g(p(x + iy)) dx dy \\ &= \frac{1}{p} \int_{p\mathcal{D}} \overline{f((x + iy)/p)} g(x + iy) dx dy \\ &= \langle f_{|A_p^{-1}}, g\rangle_\Gamma.\end{aligned}$$

Comme $\left\{ \begin{pmatrix} 1 & j \\ 0 & 1\end{pmatrix}, 0 \le j \le p - 1 \right\}$ est un système de représentants de $\Gamma_0(p)\backslash\Gamma$ et que f et $f_{|A_p^{-1}}$ sont des formes modulaires pour Γ, grâce au lemme 12 et avec les notations (2.2) et (3.1) de [AL70], on a

$$\langle f_{|A_1}, g_{|A_p}\rangle_{p^2} = \langle f_{|U_p}, g\rangle_p = -\chi(p)\langle f, g\rangle_p \tag{II.10}$$

d'après le théorème 3 de [AL70] comme f est une forme nouvelle propre de $\Gamma_0(p)$. Ainsi, V_f et V_g sont orthogonaux si f et g le sont sur $\Gamma_0(p)$ et

$$\overline{M}^t M = \overline{\left(\langle f, f\rangle_p \begin{pmatrix} p & -\chi(p) \\ -\chi(p) & p\end{pmatrix}\right)}^{-1} = \frac{1}{(p^2 - 1)\langle f, f\rangle_p} \begin{pmatrix} p & \chi(p) \\ \chi(p) & p\end{pmatrix}.$$

On obtient donc

$$\langle f,f\rangle_p(a_1,L_\chi)_{V_f}=\frac{1}{p^2-1}{}^t\begin{pmatrix}a_1(f)\\0\end{pmatrix}\begin{pmatrix}p&\chi(p)\\\chi(p)&p\end{pmatrix}\begin{pmatrix}L_\chi(f)\\\chi(p)L_\chi(f)\end{pmatrix}=\frac{a_1(f)L_\chi(f)}{p-1}.$$

En sommant ces égalités sur une base orthonormale propre de $S_2(\Gamma_0(p))^{\chi(p)}$, on obtient $(a_1,L_\chi)_{p^2}^{\mathrm{old}}$. Ainsi,

$$(a_1,L_\chi)_{p^2}^{\mathrm{old}}=\sum_f\frac{a_1(f)L_\chi(f)}{p-1}=\frac{(a_1,L_\chi)_p^{\chi(p)}}{p-1}$$

où f parcourt une base propre de $S_2(\Gamma_0(p))^{\chi(p)}$, et $(a_1,L_\chi)_p^{-\chi(p)}=0$ par l'argument du début de la preuve.

(b) La fonction L_χ est nulle sur $S_2(\Gamma_0(dp^2))^{+_p,d-\mathrm{old}}$. En effet, pour $f\in S_2(\Gamma_0(p^2))$, on a dans $S_2(\Gamma_0(dp^2))$ les égalités

$$\left(f_{|A_1}\right)_{|w_{p^2}}=\left(f_{|w_p}\right)_{|A_1},\quad \left(f_{|A_d}\right)_{|w_{p^2}}=\left(f_{|w_{p^2}}\right)_{|A_d}$$

d'après le lemme 26 [AL70], comme p et d sont premiers entre eux.

On en déduit que $S_2(\Gamma_0(dp^2))^{+_{p^2},d-\mathrm{old}}$ est engendré par les $f_{|A_1},f_{|A_d}$ avec $f\in S_2(\Gamma_0(p^2))^{+_{p^2}}$, mais alors comme dans la formule (II.7) (pas de différence entre d et p pour ceci),

$$L_\chi(f_{|A_1})=L_\chi(f)=0\quad\text{et}\quad L_\chi(f_{|A_d})=\chi(d)L_\chi(f)=0,$$

car le signe de l'équation fonctionnelle de L_χ est -1 pour un tel f. On a donc

$$(a_1,L_\chi)_{dp^2}^{+_p,-_d,\mathrm{new}}=(a_1,L_\chi)_{dp^2}^{+_p,-_d,p-\mathrm{new}}=(a_1,L_\chi)_{dp^2}^{+_p,-_d}-(a_1,L_\chi)_{dp^2}^{+_p,-_d,p-\mathrm{old}},$$

et il reste à calculer la contribution p-ancienne.

Cette contribution se décompose elle-même en deux parties : celle de $S_2(\Gamma_0(dp))^{p-\mathrm{new}}$ et celle de $S_2(\Gamma_0(d))$. Pour la seconde, comme $L_\chi(f_{|A_p})$ et $L_\chi(f_{|A_{p^2}})$ sont nuls si $L_\chi(f)$ l'est, la contribution des formes dégénérées de $S_2(\Gamma_0(d))$ à $(a_1,L_\chi)_{dp^2}$ est nulle à moins que L_χ soit déja non nulle sur $S_2(\Gamma_0(d))$, donc sur $S_2(\Gamma_0(d))^-$ vu le signe de l'équation fonctionnelle. Nous allons donc supposer cette contribution nulle. Soient f et g deux formes nouvelles propres de $S_2(\Gamma_0(dp))$. On a, par le même raisonnement que pour les formules (II.8), (II.9) et (II.10) du (a),

$$\langle f_{|A_1},g_{|A_1}\rangle_{dp^2}=\langle f_{|A_p},g_{|A_p}\rangle_{dp^2}=p\langle f,g\rangle_{dp},\quad\text{et}\quad\langle f_{|A_1},g_{|A_p}\rangle_{dp^2}=\langle f_{|U_p},g\rangle_{dp}.$$

Nous allons maintenant utiliser le fait que f est nouvelle et propre : d'après le théorème 3 de [AL70], f est propre pour U_p et W_p (vu dans $S_2(\Gamma_0(dp))$ et si $f_{|W_p}=\varepsilon f$, alors $f_{|U_p}=-\varepsilon f$. On a donc

$$\langle f_{|A_1},g_{|A_p}\rangle_{dp^2}=-\varepsilon\langle f,g\rangle_{dp}.$$

Soit maintenant $\mathcal{B}$ une base orthonormale propre de $S_2(\Gamma_0(dp))^{\mathrm{new}}$. Les espaces $\mathrm{Vect}(f_{|A_1},f_{|A_p})$ sont deux à deux orthogonaux lorsque f parcourt $\mathcal{B}$ grâce aux formules ci-dessus, et on va construire une base orthonormale de $S_2(\Gamma_0(dp))^{+_p,-_d,p-\mathrm{new}}$ respectant cette décomposition.

D'après le lemme 26 de [AL70], comme d est premier à p, pour $f\in\mathcal{B}$, $f_{|A_1}$ et $f_{|A_p}$ sont également propres pour W_d et de même valeur propre que f l'était sur $S_2(\Gamma_0(dp))$. Ainsi, la contribution de $\mathcal{B}$ à $S_2(\Gamma_0(dp))^{+_p,-_d,p-\mathrm{new}}$ se réduit aux $f\in\mathcal{B}$ tels que $f_{|w_d}=-f$ et $f_{|w_p}=f$ (on note $\mathcal{B}'$ cette partie). Il reste maintenant à produire des vecteurs propres de valeur propre 1 pour w_p. Encore une fois d'après le lemme 26 de [AL70], on a dans $S_2(\Gamma_0(dp^2))$

$$(f_{|w_p})_{A_p}=(f_{|A_1})_{|w_{p^2}},\qquad(f_{|A_p})_{|w_{p^2}}=(f_{|w_p})_{|A_1}.$$

Une base orthogonale de $S_2(\Gamma_0(dp))^{+_p,-_d,p-\mathrm{new}}$ est donc constituée des

$$f_{|A_1}+\left(f_{|A_1}\right)_{|w_{p^2}}=f_{|A_1}\pm f_{|A_p},\quad f\in\mathcal{B}'.$$

Or, d'après les formules précédentes, pour $\varepsilon = \pm 1$ et $f \in \mathcal{B}'$:

$$\langle f_{|A_1} + \varepsilon f_{|A_p}, f_{|A_1} + \varepsilon f_{|A_p} \rangle_{dp^2} = 2p - 2\varepsilon^2 = 2(p-1).$$

Enfin,

$$\overline{a_1}(f_{|A_1} + \varepsilon f_{|A_p}) L_\chi(f_{|A_1} + \varepsilon f_{|A_p}) = \overline{a_1(f)}(L_\chi(f) + \varepsilon\chi(p) L_\chi(f)).$$

Regroupant tous ces termes et après orthonormalisation, on obtient donc

$$(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,p-\mathrm{old}} = \frac{1}{p-1}(a_1, L_\chi)_{dp}^{\chi(p)_p,-_d}.$$

□

Il reste désormais à évaluer précisément les termes du lemme II.1.1.

Pour ceci, nous allons employer la formule des traces de Petersson classique ([IK04], Proposition 14.5) et une formule « restreinte », que nous énonçons et démontrons dans la section suivante.

II.2 Formules des traces de Petersson et Akbary

Cette section peut se retrouver (en anglais) dans [LF]. Commençons par les définitions nécessaires pour les formules des traces.

Définition II.2.1 (Sommes de Kloosterman et fonction de Bessel)**.**

Pour trois entiers m, n, c avec $c > 0$, la *somme de Kloosterman associée à* m, n, c est la somme définie par

$$S(m, n; c) = \sum_{k \in (\mathbb{Z}/c\mathbb{Z})^*} e^{2i\pi(mk + nk^{-1})/c}.$$

Ces sommes admettent les bornes dites de Weil ([IK04], Corollaire 11.12)

$$|S(m, n; c)| \leq (m, n, c)^{1/2} \tau(c) \sqrt{c}, \tag{II.11}$$

avec (m, n, c) le pgcd de m,n et c et $\tau(c)$ le nombre de diviseurs positifs de c. En fait, pour $c = p^\alpha c'$ avec p premier impair et $(p, c') = 1$, on peut plus précisément écrire la borne

$$|S(m, n; c)| \leq 2\tau(c')(m, n, c)^{1/2} \sqrt{c} \tag{II.12}$$

grâce aux calculs de Salié ([IK04], formule 12.39).

La *fonction de Bessel de premier type d'ordre 1* est la fonction entière J_1 définie par la série entière de rayon de convergence infini

$$J_1(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!(n+1)!} \left(\frac{z}{2}\right)^{2n+1}.$$

Cette fonction admet la représentation intégrale suivante ([Wat22], 6.21, Formule 8)

$$J_1(z) = \frac{z}{4i\pi} \int_{x-i\infty}^{x+i\infty} \frac{e^{w - \frac{z^2}{4w}}}{w^2} dw, \tag{II.13}$$

pour tout $z \in \mathbb{C}$ et tout réel x strictement positif.

L'objectif de cette sous-section est de prouver les formules suivantes.

Proposition II.2.2 (Formules des traces)**.**

Soient m, n, N trois entiers plus grands que 1. Alors,

$$\frac{1}{4\pi\sqrt{mn}}(a_m, a_n)_N = \delta_{mn} - 2\pi \sum_{\substack{c>0 \\ N|c}} \frac{S(m,n;c)}{c} J_1\left(\frac{4\pi\sqrt{mn}}{c}\right). \tag{II.14}$$

Avec les notations du paragraphe précédent, pour tout diviseur positif Q de N avec $Q > 1$, on a également pour $\varepsilon = \pm 1$:

$$\begin{aligned}\frac{1}{2\pi\sqrt{mn}}(a_m, a_n)_N^{\varepsilon Q} = \delta_{mn} \quad &- \quad 2\pi \sum_{\substack{c>0 \\ N|c}} \frac{S(m,n;c)}{c} J_1\left(\frac{4\pi\sqrt{mn}}{c}\right) \\ &- \quad 2\pi\varepsilon \sum_{\substack{c>0 \\ (N/Q)|c \\ (c,Q)=1}} \frac{S(m,nQ^{-1};c)}{c\sqrt{Q}} J_1\left(\frac{4\pi\sqrt{mn}}{c\sqrt{Q}}\right).,\end{aligned} \tag{II.15}$$

où la notation nQ^{-1} dans la somme de Kloosterman désigne n multiplié par l'inverse de Q modulo c.

Remarque II.2.1. La formule (II.14) est la formule des traces de Petersson originelle (proposition 14.5 de [IK04]), et se généralise à tout poids $k \geq 2$. Cependant, la preuve pour $k = 2$ mobilise par défaut de convergence uniforme une définition différente des séries de Poincaré, qui ne semble être complètement traitée que dans la section 5.5 de [Ran77] (et pour $\Gamma(N)$ au lieu de $\Gamma_0(N)$, bien que les idées soient exactement les mêmes). Nous allons donc ci-dessous reprouver en détail ces formules pour le confort du lecteur, sachant que nombre d'outils nous serviront ensuite également pour (II.15). A noter que cette dernière formule (pour $Q = N$ seulement mais tout poids $k \geq 2$) est l'objet du chapitre 3 de [Akb97].

Enfin, on peut remarquer que la formule des traces restreinte montre que le terme d'erreur du corollaire 14.26 de [IK04] est en fait nul, ce qui pourrait éventuellement améliorer les estimations de ce livre qui en découlent.

Pour prouver ces formules des traces, nous allons utiliser les séries de Poincaré de poids 2, dont les propriétés seront prouvées ci-dessous.

Définition-Proposition II.2.3 (Séries de Poincaré de poids 2)**.**

Pour tous entiers $N, n \geq 1$, il existe des formes cuspidales de poids 2 pour $\Gamma_0(N)$ notées $P_n(\cdot, N)$ et appelées *séries de Poincaré de poids 2* telles que :

(a) Pour tout $m \in \mathbb{N}$,

$$a_m(P_n(\cdot, N)) = \delta_{mn} - 2\pi\left(\sqrt{\frac{m}{n}} \sum_{\substack{c>0 \\ N|c}} \frac{S(m,n;c)}{c} J_1\left(\frac{4\pi\sqrt{mn}}{c}\right)\right). \tag{II.16}$$

(b) Pour tout diviseur $Q > 1$ de N tel que $(Q, N/Q) = 1$ et tout $m \geq 1$,

$$a_m(P_n(\cdot, N)_{|w_Q}) = -2\pi\sqrt{\frac{m}{n}} \sum_{m\geq 1} \sum_{\substack{c>0 \\ (N/Q)|c \\ (Q,c)=1}} \frac{S(m,nQ^{-1};c)}{c\sqrt{Q}} J_1\left(\frac{4\pi\sqrt{mn}}{c\sqrt{Q}}\right). \tag{II.17}$$

(c) Pour toute forme modulaire $f \in S_2(\Gamma_0(N))$, $\langle f, P_n(\cdot, N)\rangle_N = \overline{a_n(f)}/(4\pi n)$.

Remarque II.2.2. Les coefficients de Fourier donnés dans le (b) permettent des formules des traces restreintes même en combinant plusieurs contraintes de signes pour des involutions d'Atkin-Lehner distinctes : par exemple, pour Q et Q' premiers entre eux, on procède comme dans la preuve ci-dessous avec $P_n(\cdot, N) + P_n(\cdot, N)_{|w_Q} + P_n(\cdot, N)_{|w_{Q'}} + P_n(\cdot, N)_{|w_{QQ'}}$ pour obtenir la formule des traces sur $S_2(\Gamma_0(N))^{+_Q,+_{Q'}}$ et on obtient les autres combinaisons possibles par différence.

Démonstration. Commençons par expliquer pourquoi les propriétés (a), (b) et (c) des séries de Poincaré de poids 2 impliquent les formules des traces.

Soit $\mathcal{F}_N$ une base orthonormale de $S_2(\Gamma_0(N))$. D'après le (c), pour tout $n \in \mathbb{N}$,

$$P_n = \sum_{f \in \mathcal{F}_N} \langle f, P_n(\cdot, N)\rangle f = \sum_{f \in \mathcal{F}_N} \frac{\overline{a_n(f)}}{4\pi n} f.$$

En prenant le m-ième coefficient de Fourier des deux cotés, on obtient alors d'après le (a)

$$\sum_{f \in \mathcal{F}_N} \frac{a_m(f)\overline{a_n(f)}}{4\pi n} = \delta_{mn} - 2\pi\sqrt{\frac{m}{n}} \sum_{\substack{c>0 \\ N|c}} \frac{S(m,n;c)}{c} J_1\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

Le terme de droite est réel (car les sommes de Kloosterman le sont), donc après conjugaison complexe et multiplication par $\sqrt{n/m}$ on obtient la formule des traces de Petersson, à savoir

$$\sum_{f \in \mathcal{F}_N} \frac{\overline{a_m(f)}a_n(f)}{4\pi\sqrt{mn}} = \delta_{mn} - 2\pi \sum_{\substack{c>0 \\ N|c}} \frac{S(m,n;c)}{c} J_1\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

Pour $Q > 1$ divisant N tel que $(Q, N/Q) = 1$, rappelons que l'opérateur w_Q est auto-adjoint pour le produit scalaire de Petersson, et définissons

$$P_n^{+Q}(\cdot, N) = P_n(\cdot, N) + P_n(\cdot, N)_{|w_Q}$$

C'est bien une forme modulaire cuspidale de poids 2 pour $\Gamma_0(N)$, invariante par w_Q et pour toute forme modulaire $f \in S_2(\Gamma_0(N))^{+Q}$, on a alors

$$\langle f, P_n^{+Q}(\cdot, N)\rangle = \langle f, P_n(\cdot, N)\rangle + \langle f, P_n(\cdot, N)_{|w_Q}\rangle = \langle f, P_n(\cdot, N)\rangle + \langle f_{|w_Q}, P_n(\cdot, N)\rangle = 2\langle f, P_n(\cdot, N)\rangle.$$

On en déduit que pour f parcourant une base orthonormale $\mathcal{F}_{N,Q}$ de $S_2(\Gamma_0(N))^{+Q}$, d'après le (c),

$$P_n^{+Q}(\cdot, N) = \sum_{f \in \mathcal{F}_{N,Q}} \langle f, P_n^{+Q}\rangle f = 2 \sum_{f \in \mathcal{F}_{N,Q}} \langle f, P_n\rangle f = \sum_{f \in \mathcal{F}_{N,Q}} \frac{\overline{a_n(f)}}{2\pi n} f.$$

Le (a) et le (b) nous donnent les coefficients de Fourier de $P_n^{+Q}(\cdot, N)$, donc la formule des traces pour Q et $\varepsilon = 1$ par le même raisonnement que précédemment. On en déduit celle pour $\varepsilon = -1$ par différence entre les deux déja obtenues.

Nous allons maintenant construire les séries de Poincaré de poids 2 et prouver dans l'ordre (a), (b) et (c).

Soit $\Gamma_\infty = \left\{ \pm \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, k \in \mathbb{Z} \right\}$ le groupe parabolique. Pour $\gamma, \gamma' \in \mathrm{GL}_2^+(\mathbb{R})$, si $\Gamma_\infty\gamma = \Gamma_\infty\gamma'$ alors γ et γ' ont au signe près la même deuxième ligne. En conséquence, si $\overline{\gamma}$ est une classe non triviale de $\Gamma_\infty\backslash\Gamma_0(N)$ avec $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, le coefficient c est non nul et il existe alors un unique représentant γ de $\overline{\gamma}$ tel que $0 \leq a < c$. L'ensemble des classes non triviales de $\Gamma_\infty\backslash\Gamma_0(N)$ correspond donc canoniquement à $\mathcal{R}_N$ l'ensemble des triplets d'entiers (a, c, d) tels que N divise c, $1 \leq a \leq c$ et $ad = 1 \mod c$. Ensuite, pour tout $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, on a

$$\left| \frac{e^{2i\pi n\gamma\cdot z}}{j_\gamma(z)^2 |j_\gamma(z)|^{2s}} \right| = \frac{e^{-\frac{2\pi n \operatorname{Im} z}{|j_\gamma(z)|^2}}}{|j_\gamma(z)|^{2+2s}} \leq \frac{1}{(c^2(\operatorname{Im} z)^2 + d^2)^{1+s}}.$$

Or, pour tout $\eta > 0$,

$$\sum_{(a,c,d) \in \mathcal{R}_N} \frac{1}{(c^2\eta^2 + d^2)^{1+s}} \leq \sum_{\substack{(c,d) \in \mathbb{Z}^2 \\ (c,d) \neq (0,0)}} \frac{1}{(c^2\eta^2 + d^2)^{1+s}} < +\infty$$

car $s > 0$. Ainsi, la somme

$$P_n(z,s,N) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \frac{e^{2i\pi n \gamma\cdot z}}{j_\gamma(z)^2 |j_\gamma(z)|^{2s}}$$

(qui ne dépend pas des choix de représentants) converge normalement sur tout domaine $\mathrm{Im}(z) \geq \eta$ pour $\eta > 0$ et définit donc une fonction continue sur $\mathcal{H}$. De plus, la convergence étant uniforme, pour tout $\gamma \in \Gamma_0(N)$,

$$P_n(\gamma \cdot z, s, N) = j_\gamma(z)^2 |j_\gamma(z)|^{2s} P_n(z,s,N) \tag{II.18}$$

par la propriété de cocycle de j.

Nous allons maintenant montrer que les $P_n(\cdot, s, N)$ convergent uniformément sur tout compact lorsque $s \to 0^+$, en les développant en série de Fourier. On en déduira simultanément la définition de $P_n(\cdot, N)$ comme cette limite et son développement en série de Fourier.

Notons $\mathcal{R}'_N$ l'ensemble des triplets (a,c,d) de $\mathcal{R}_N$ tels que $0 \leq d < c$. On peut alors écrire

$$\begin{aligned} P_n(z,s,N) &= e^{2i\pi nz} + \sum_{(a,c,d)\in\mathcal{R}_N} \frac{e^{2i\pi n\left(\frac{a}{c} - \frac{1}{c(cz+d)}\right)}}{(cz+d)^2|cz+d|^{2s}} \\ &= e^{2i\pi nz} + \sum_{(a,c,d)\in\mathcal{R}'_N} \frac{e^{2i\pi na/c}}{c^{2+2s}} \sum_{\ell\in\mathbb{Z}} \frac{e^{-\frac{2i\pi n}{c^2(z+d/c+\ell)}}}{(z+d/c+\ell)^2|z+d/c+\ell|^{2s}} \\ &= e^{2i\pi nz} + \sum_{(a,c,d)\in\mathcal{R}'_N} \frac{e^{2i\pi na/c}}{c^{2+2s}} F_{c,n,s}(z+d/c) \end{aligned}$$

où

$$F_{c,n,s}(z) := \sum_{\ell\in\mathbb{Z}} f_{c,n,s,z}(\ell) \quad \text{avec} \quad f_{c,n,s,z}(x) := \frac{e^{-\frac{2i\pi n}{c^2(x+z)}}}{(x+z)^2|x+z|^{2s}}\ . \tag{II.19}$$

Or, la fonction $f_{c,n,s,z}$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, et on vérifie immédiatement qu'elle est intégrable ainsi que sa dérivée et sa dérivée seconde. On peut donc appliquer la formule sommatoire de Poisson pour réécrire

$$F_{c,n,s}(z) = \sum_{m\in\mathbb{Z}} \int_{-\infty}^{+\infty} f_{c,n,s,z}(x) e^{-2i\pi mx} dx.$$

On fixe désormais $\eta > 0$, et on se restreint au domaine $\mathrm{Im}\, z \geq \eta$. Alors, la fonction $f_{c,n,s,z}$ s'étend en une fonction holomorphe sur $|\mathrm{Im}\, x| < \eta$ lorsqu'on utilise la détermination usuelle du logarithme sur $\mathbb{C}\backslash\mathbb{R}^-$ pour écrire, pour tout $x \in \mathbb{R}$,

$$(x+z)^2|x+z|^{2s} = (x+z)^2(x+z)^s(x+\overline{z})^s.$$

On prolonge donc cette fonction holomorphe en x (z étant fixé) sur le domaine $|\mathrm{Im}\, x| < \eta$. Par définition de la détermination du logarithme, $|(x+\overline{z})^s| = |x+\overline{z}|^s$. Comme l'intégrande est holomorphe sur ce domaine, on peut décaler la partie imaginaire d'intégration de $\varepsilon\eta/2$, avec $\varepsilon = -1$ si $m > 0$ et $\varepsilon = 1$ si $m < 0$, et alors $\mathrm{Re}(-2i\pi mx) = 2\pi m\, \mathrm{Im}(x) = -\pi|m|\eta$, d'où

$$\begin{aligned} \left| \int_{-\infty}^{+\infty} f_{c,n,s,z}(x) e^{-2i\pi mx} dx \right| &= \left| \int_{i\varepsilon\eta/2+\mathbb{R}} f_{c,n,s,z}(x) e^{-2i\pi mx} dx \right| \\ &\leq \int_{i\varepsilon\eta/2+\mathbb{R}} \frac{e^{-\pi|m|\eta}}{|x+z|^{2+2s}} dx \\ &\leq e^{-\pi|m|\eta} \int_{\mathbb{R}} \frac{1}{(\eta^2/4+x^2)^{1+s}} dx. \end{aligned}$$

De plus, par translation réelle dans l'intégrale, pour tout réel y et tout entier $m \in \mathbb{Z}$,

$$\int_{\mathbb{R}} f_{c,n,s,z+y}(x)e^{-2i\pi m(x+y)}dx = \int_{\mathbb{R}} f_{c,n,s,z}(x)e^{-2i\pi mx}dx.$$

On peut donc écrire, en notant $\mathcal{R}_c = \{(a,d) \in \mathbb{N} \mid a,d \leq c,\ ad = 1 \mod c\}$,

$$\begin{aligned} P_n(z,s,N) &= e^{2i\pi nz} + \sum_{\substack{c>0\\ N|c}} \sum_{(a,d)\in\mathcal{R}_c} \frac{e^{2i\pi na/c}}{c^{2+2s}} \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}} f_{c,n,s,z+d/c}(x)e^{-2i\pi mx}dx \\ &= e^{2i\pi nz} + \sum_{\substack{c>0\\ N|c}} \sum_{(a,d)\in\mathcal{R}_c} \sum_{m\in\mathbb{Z}} \frac{e^{2i\pi(md+na)/c}}{c^{2+2s}} \int_{\mathbb{R}} f_{c,n,s,z}(x)e^{-2i\pi mx}dx \\ &= e^{2i\pi nz} + \sum_{\substack{c>0\\ N|c}} \sum_{m\in\mathbb{Z}} \frac{S(m,n;c)}{c^{2+2s}} \int_{\mathbb{R}} f_{c,n,s,z}(x)e^{-2i\pi mx}dx. \end{aligned}$$

D'après les bornes de Weil (II.11), on a

$$\left| \frac{S(m,n;c)}{c^{2+2s}} \int_{\mathbb{R}} |f_{c,n,s,z}(x)e^{-2i\pi mx}|dx \right| \leq \frac{n^{1/2}\tau(c)}{c^{3/2}} e^{-\pi|m|\eta} \int_{\mathbb{R}} \frac{1}{\eta^2/4 + x^2}dx,$$

qui est le terme général d'une série absolument convergente (indépendante de s) car $\tau(c) = O(c^\varepsilon)$ pour tout $\varepsilon > 0$. On peut donc permuter la somme double, et on obtient ainsi

$$P_n(z,s,N) = e^{2i\pi nz} + \sum_{m\in\mathbb{Z}} \left(\sum_{\substack{c>0\\ N|c}} \frac{S(m,n;c)}{c^{2+2s}} \int_{\mathbb{R}} f_{c,n,s,z}(x)e^{-2i\pi m(x+z)}dx \right) e^{2i\pi mz}.$$

De plus, la convergence étant uniforme sur $\operatorname{Im} z \geq \eta$ pour tout $\eta > 0$ comme on vient de le montrer et indépendante de s, on peut faire tendre s vers 0 par convergence dominée. On obtient comme limite simple la série de Poincaré de poids 2

$$P_n(z,N) := \lim_{s\to 0^+} P_n(z,s,N) = e^{2i\pi nz} + \sum_{m\in\mathbb{Z}} \left(\sum_{\substack{c>0\\ N|c}} \frac{S(m,n;c)}{c^2} \int_{\mathbb{R}} \frac{e^{-\frac{2i\pi n}{c^2(x+z)} - 2i\pi m(x+z)}}{(x+z)^2}dx \right) e^{2i\pi mz}.$$

Il nous reste maintenant à calculer ces intégrales. Posons

$$G_{m,n,c}(z) := \int_{\mathbb{R}} \frac{e^{-\frac{2i\pi n}{c^2(x+z)} - 2i\pi m(x+z)}}{(x+z)^2}dx = \int_{i\operatorname{Im}(z)+\mathbb{R}} \frac{e^{-\frac{2i\pi n}{c^2 y} - 2i\pi my}}{y^2}dy. \tag{II.20}$$

L'intégrande est ici une fonction holomorphe sur $\mathbb{C}^*$, donc par le théorème des résidus, on peut intégrer sur n'importe quelle ligne horizontale d'ordonnée strictement positive α, ainsi $G_{m,n,c}(z)$ ne dépend pas de z, on le renote $G_{m,n,c}$. Pour $m \leq 0$,

$$|G_{m,n,c}| = \left| \int_{i\alpha+\mathbb{R}} \frac{e^{-\frac{2i\pi n}{c^2 y} - 2i\pi my}}{y^2}dy \right| \leq \int_{i\alpha+\mathbb{R}} \frac{e^{2\pi m\alpha}}{|y|^2}dy$$

et ceci tend vers 0 quand α tend vers $+\infty$, donc $G_{m,n,c} = 0$ lorsque $m \leq 0$.

Maintenant, pour $m > 0$,

$$\begin{aligned} G_{m,n,c} &= \int_{i+\mathbb{R}} \frac{e^{-\frac{2i\pi n}{c^2 y} - 2i\pi my}}{y^2}dy \\ &= 2i\pi m \int_{2\pi m - i\infty}^{2\pi m + i\infty} \frac{e^{w - \frac{4\pi^2 mn}{c^2 w}}}{w^2}dw, \qquad w = -2i\pi m \\ &= 2i\pi m J_1\left(\frac{4\pi\sqrt{mn}}{c}\right) \frac{ic}{\sqrt{mn}} \end{aligned}$$

grâce à la représentation intégrale (II.13) de la fonction J_1. On obtient finalement

$$G_{m,n,c} = -2\pi c\sqrt{\frac{m}{n}}J_1\left(\frac{4\pi\sqrt{mn}}{c}\right), \tag{II.21}$$

d'où le développement en série de Fourier

$$P_n(z,N) = e^{2i\pi nz} - 2\pi\sum_{m\geq 1}\left(\sqrt{\frac{m}{n}}\sum_{\substack{c>0\\N|c}}\frac{S(m,n;c)}{c}J_1\left(\frac{4\pi\sqrt{mn}}{c}\right)\right)e^{2i\pi mz}.$$

En passant à la limite dans la formule (II.18), on voit que $P_n(\cdot,N)$ est une forme modulaire de poids 2 pour $\Gamma_0(N)$, et elle est de plus cuspidale (on peut montrer que $P_n(\cdot,s,N)$ tend uniformément en s vers 0 en chaque pointe de $\Gamma_0(N)$), ce qui conclut la preuve du (a).

La preuve du (b) est assez similaire. Pour tout $s>0$, avec un choix de matrice W_Q, on a

$$\begin{aligned}
P_n(\cdot,s,N)_{|W_Q}(z) &= \frac{\det W_Q}{j_{W_Q}(z)^2}P_n(W_Q\cdot z,s,N)\\
&= \frac{Q}{j_{W_Q}(z)^2}\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)}\frac{e^{2i\pi n\gamma W_Q z}}{j_\gamma(W_Q z)^2|j_\gamma(W_Q z)|^{2s}}\\
&= Q|j_{W_Q}(z)|^{2s}\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)W_Q}\frac{e^{2i\pi n\gamma z}}{j_\gamma(z)^2|j_\gamma(z)|^{2s}}
\end{aligned}$$

par la relation de cocycle de j. Pour toute matrice $\begin{pmatrix}a&b\\c&d\end{pmatrix}\in\Gamma_0(N)$,

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}W_Q = \begin{pmatrix}aQ+bN & ay+bQt\\ cQ+dN & cy+dQt\end{pmatrix}$$

et appartient donc à l'ensemble des matrices $\begin{pmatrix}a'&b'\\c'&d'\end{pmatrix}$ à coefficients entiers tels que N divise c', Q divise a' et d', et de déterminant Q. En fait, cet ensemble est exactement $\Gamma_0(N)W_Q$ comme on le vérifie immédiatement par multiplication par W_Q^{-1}, et pour $Q>1$, c' est forcément non nul, donc $\Gamma_\infty\backslash\Gamma_0(N)W_Q$ est en bijection naturelle avec l'ensemble $\mathcal{R}_{N,Q}$ des triplets (a,c,d) d'entiers tels que $c>0, N|c$, $Q|(a,d)$, $ad=Q \mod c$ et $0\leq a<c$ par un raisonnement similaire au (a). Pour $\gamma=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ construit dans $\Gamma_0(N)W_Q$ à partir d'une telle classe, on a d'après (II.1)

$$\gamma\cdot z = \frac{a}{c} - \frac{Q}{c(cz+d)},$$

d'où

$$\begin{aligned}
P_n(\cdot,s,N)_{|W_Q}(z) &= Q|j_{W_Q}(z)|^{2s}\sum_{(a,c,d)\in\mathcal{R}_{N,Q}}\frac{e^{2i\pi na/c}e^{-\frac{2i\pi nQ}{c(cz+d)}}}{(cz+d)^2|cz+d|^{2s}}\\
&= Q|j_{W_Q}(z)|^{2s}\sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}}\frac{1}{c^{2+2s}}\sum_{\substack{0\leq a<c\\Q|a}}e^{2i\pi na/c}\sum_{\substack{Q|d\\ad\equiv Q[c]}}\frac{e^{-\frac{2i\pi nQ}{c^2(z+d/c)}}}{(z+d/c)^2|z+d/c|^{2s}}.
\end{aligned}$$

Pour a et c fixés, l'ensemble des d vérifiant $Q|d$ et $ad\equiv Q[c]$ est une classe de congruence modulo c, dont on choisit le représentant d' entre 0 et c. On a alors

$$\sum_{\substack{Q|d\\ad\equiv Q[c]}}\frac{e^{-\frac{2i\pi nQ}{c^2(z+d/c)}}}{(z+d/c)^2|z+d/c|^{2s}} = \sum_{\ell\in\mathbb{Z}}\frac{e^{-\frac{2i\pi nQ}{c^2(z+d'/c+\ell)}}}{(z+d'/c+\ell)^2|z+d'/c+\ell|^{2s}} = F_{c/\sqrt{Q},n,s}(z+d'/c)$$

avec les notations (II.19). On réutilise le même raisonnement (passant par la formule sommatoire de Poisson), de sorte que

$$\begin{aligned}\frac{P_n(\cdot,s,N)_{|W_Q}(z)}{Q|j_{W_Q}(z)|^{2s}} &= \sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}} \frac{1}{c^{2+2s}} \sum_{\substack{0\le a,d\le c\\Q|(a,d)\\ad\equiv Q[c]}} e^{2i\pi na/c} \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}} f_{c/\sqrt{Q},n,s,z+d/c}(x) e^{-2i\pi mx}dx\\ &= \sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}} \frac{1}{c^{2+2s}} \sum_{\substack{0\le a,d\le c\\Q|(a,d)\\ad\equiv Q[c]}} e^{2i\pi (na+md)/c} \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}} f_{c/\sqrt{Q},n,s,z}(x) e^{-2i\pi mx}dx\\ &= \sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}} \frac{1}{c^{2+2s}} \sum_{m\in\mathbb{Z}} \sum_{\substack{0\le a,d\le c\\Q|(a,d)\\ad\equiv Q[c]}} e^{2i\pi (na+md)/c} \int_{\mathbb{R}} f_{c/\sqrt{Q},n,s,z}(x) e^{-2i\pi mx}dx.\end{aligned}$$

Ici, pour c fixé, a et d parcourent les multiples de Q entre 0 et c tels que $ad = Q \mod c$. Cela revient à dire que $a = Qa'$ et $d = Qd'$ avec a', d' entre 0 et c/Q tels que $Qa'd' = 1 \mod c$, c'est-à-dire que d' est égal à $Q^{-1}a'^{-1}$ modulo c/Q. On en déduit l'égalité

$$\sum_{\substack{0\le a,d\le c\\Q|(a,d)\\ad\equiv Q[c]}} e^{2i\pi (na+md)/c} = S(m, nQ^{-1}; c/Q)$$

où Q^{-1} est l'inverse de Q modulo c/Q. En conséquence,

$$P_n(\cdot,s,N)_{|W_Q}(z) = Q|j_{W_Q}(z)|^{2s} \sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}} \frac{S(m,nQ^{-1};c/Q)}{c^{2+2s}} \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}} f_{c/\sqrt{Q},n,s,z}(x) e^{-2i\pi mx}dx.$$

De manière parfaitement similaire à la preuve du (a), grâce aux bornes de Weil (II.11) la permutation somme-intégrale est valide, et on peut faire tendre s vers 0 par convergence dominée pour obtenir

$$P_n(\cdot,N)_{|W_Q}(z) = Q \sum_{m\in\mathbb{Z}} \sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}} \frac{S(m,nQ^{-1};c/Q)}{c^2} G_{m,n,c/\sqrt{Q}} e^{2i\pi mz}$$

avec $G_{m,n,c}$ défini dans (II.20) et calculé dans (II.21), ce qui nous donne

$$\begin{aligned}P_n(\cdot,N)_{|W_Q}(z) &= -2\pi Q\sqrt{m/n} \sum_{m\ge 1} \sum_{\substack{c>0\\N|c\\(Q,c/Q)=1}} \frac{S(m,nQ^{-1};c/Q)}{c^2} c/\sqrt{Q} J_1\left(\frac{4\pi\sqrt{mn}}{c/\sqrt{Q}}\right) e^{2i\pi mz}\\ &= -2\pi\sqrt{m/n} \sum_{m\ge 1} \sum_{\substack{c>0\\(N/Q)|c\\(Q,c)=1}} \frac{S(m,nQ^{-1};c)}{c\sqrt{Q}} J_1\left(\frac{4\pi\sqrt{mn}}{c\sqrt{Q}}\right) e^{2i\pi mz}\end{aligned}$$

après réindexation de c par c/Q, ce qui conclut la preuve du (b).

Enfin, montrons le (c). Soit $f \in S_2(\Gamma_0(N))$, $s > 0$ et D un domaine fondamental de $\mathcal{H}$ pour $\Gamma_0(N)$. Comme Γ_∞ admet pour domaine fondamental $D' = \{z \in \mathcal{H}, -1/2 \le \operatorname{Re} z \le 1/2\}$, on peut choisir D et les représentants γ de $\Gamma_\infty\backslash\Gamma_0(N)$ de sorte que $D' = \bigcup_\gamma \gamma \cdot D$.

On va définir une intégrale en fonction de f et $P_n(\cdot,s,N)$, montrer qu'elle converge vers $\langle f, P_n\rangle_N$

lorsque s tend vers 0 et calculer cette limite. On pose donc

$$\begin{aligned} I(f,n,s) &:= \int_D \overline{f(z)} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \frac{e^{2i\pi n\gamma\cdot z}}{j_\gamma(z)^2|j_\gamma(z)|^{2s}} y^s dxdy \\ &= \int_D \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \overline{f(\gamma z)} \frac{e^{2i\pi n\gamma\cdot z}}{|j_\gamma(z)|^{4+2s}} y^s dxdy \\ &= \int_D \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \overline{f(\gamma z)} e^{2i\pi n\gamma\cdot z} \operatorname{Im}(\gamma z)^{2+s} \frac{dx\,dy}{y^2}. \end{aligned}$$

Cette intégrale double converge absolument : en effet,

$$\int_D \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \left| f(\gamma z) e^{2i\pi n\gamma\cdot z} \operatorname{Im}(\gamma z)^{2+s} \right| \frac{dx\,dy}{y^2} = \int_{D'} |f(z) e^{-2\pi n y} y^s| dxdy$$

et ceci converge. En effet, comme f est cuspidale, il suffit de montrer que f est intégrable sur D' pour $dxdy$. On note D_0 le domaine fondamental usuel de $\mathrm{SL}_2(\mathbb{Z})$, et $\mathcal{R}$ l'ensemble des représentants de $\Gamma_\infty\backslash\mathrm{SL}_2(\mathbb{Z})$ tel que $D' = \cup_{\gamma\in\mathcal{R}}\gamma D_0$. On a alors

$$\begin{aligned} \int_{D'} |f(z)| dxdy &= \sum_{\gamma\in\mathcal{R}} \int_{\gamma\cdot D_0} |f(z)| dxdy \\ &= \sum_{\gamma\in\mathcal{R}} \int_{\gamma\cdot D_0} |f(z)| \operatorname{Im}(z)^2 \frac{dxdy}{y^2} \\ &= \sum_{\gamma\in\mathcal{R}} \int_{D_0} |f_{|\gamma^{-1}}(z) j_{\gamma^{-1}}(z)^2 \operatorname{Im}(\gamma^{-1}z)^2 \frac{dxdy}{y^2} \\ &= \sum_{\gamma\in\mathcal{R}} \int_{D_0} \frac{f_{|\gamma^{-1}}(z)}{j_{\gamma^{-1}}(z)^2} dxdy \\ &= \sum_{\gamma\in\mathcal{R}} \int_{D_0} \frac{f_{|\gamma^{-1}}(z)}{3c^2/4 + a^2} dxdy \end{aligned}$$

or deux éléments de $\mathcal{R}$ distincts n'ont jamais la même première colonne, et $f_{|\gamma^{-1}}$ ne dépend que de $\gamma\Gamma_0(N)$ et pour chacune des classes $\gamma\Gamma_0(N)$ est intégrable sur D_0 pour $dxdy$, d'intégrale bornée par C quel que soit γ. On a donc

$$\int_{D'} |f(z)| dxdy \leq C \sum_{\substack{(a,c)\in\mathbb{Z}^2 \\ (a,c)\neq(0,0)}} \frac{1}{a^2 + 3c^2/4} < +\infty.$$

En revenant à l'intégrande de $I(f,n,s)$, on peut donc dominer celui-ci indépendamment de s (par exemple en dominant y^s par 1 pour $y < 1$ et par y pour $0 < s < 1$, ce qui nous suffira), et permuter somme et intégrale par convergence absolue. La convergence dominée assure que $I(f,n,s)$ tend

vers $\langle f, P_n(\cdot, N)\rangle_N$ lorsque s tend vers 0, et grâce aux arguments précédents

$$\begin{aligned} I(f,n,s) &= \int_{D'} \overline{f(z)} e^{2i\pi n z} y^s dxdy \\ &= \int_{-1/2}^{1/2}\int_0^{+\infty} \overline{f(x+iy)} e^{2i\pi n(x+iy)} y^s dxdy \\ &= \int_{-1/2}^{1/2}\int_0^{+\infty} \sum_{m\in\mathbb{N}^*} \overline{a_m(f)} e^{2i\pi(n-m)x - 2\pi(m+n)y} y^s dxdy \\ &= \int_0^{+\infty} y^s e^{-2\pi(m+n)y} \sum_{m\in\mathbb{N}^*} \int_{-1/2}^{1/2} \overline{a_m(f)} e^{2i\pi(n-m)x} dxdy \\ &= \overline{a_n(f)} \int_0^{+\infty} y^s e^{-4\pi n y} dy. \end{aligned}$$

Par convergence dominée, ceci tend vers $a_n(f)/(4\pi n)$ lorsque s tend vers 0^+, ce qui conclut la preuve du (c). □

II.3 Calculs de la moyenne pondérée (cas quadratique imaginaire)

On suppose ici que χ est un caractère quadratique impair de conducteur D. Le but de cette section est de démontrer le théorème II.1 (qui est également l'objet de l'Annexe de [LF]), nous donnant la non-annulation de $(a_1, L_\chi)_{p^2}^{+,p-\mathrm{new}}$ pour p assez grand.

Avec les notations des deux sections précédentes et grâce au lemme II.1.1 (a), pour p premier ne divisant pas D, nous avons besoin d'évaluer les sommes $(a_1, L_\chi)_N$ pour $N = p$ et $N = p^2$. On pose dans tous les calculs de cette section

$$x = \frac{2\pi}{D\sqrt{N}},$$

de sorte que d'après les équations fonctionnelles (II.2) et (II.3), pour tout $f \in S_2(\Gamma_0(N))^{\chi(N)}$,

$$L_\chi(f) = 4\pi \int_{1/(D\sqrt{N})}^{+\infty} (f\otimes\chi)(iu)du = 2\sum_{n=1}^{+\infty} \frac{\chi(n)a_n(f)}{n} e^{-nx}.$$

Grâce la formule restreinte (II.15) appliquée à $m = 1$ et $Q = N$, on en déduit que

$$(a_1, L_\chi)_N = (a_1, L_\chi)^{\chi(N)} = 4\pi e^{-x} - 8\pi^2(A(\chi,N) + \chi(N)B(\chi,N)) \tag{II.22}$$

où

$$\begin{aligned} A(\chi,N) &:= \sum_{n=1}^{+\infty} \frac{\chi(n)}{\sqrt{n}} e^{-nx} \sum_{\substack{c>0\\ N|c}} \frac{S(1,n;c)}{c} J_1\left(\frac{4\pi\sqrt{n}}{c}\right) \\ B(\chi,N) &:= \sum_{n=1}^{+\infty} \frac{\chi(n)}{\sqrt{n}} e^{-nx} \sum_{\substack{d>0\\ (N,d)=1}} \frac{S(1,n/N;d)}{d\sqrt{N}} J_1\left(\frac{4\pi\sqrt{n}}{d\sqrt{N}}\right). \end{aligned}$$

Avec les bornes de Weil (II.11), on peut permuter les sommes ci-dessus, d'où les réécritures

$$A(\chi,N) = \sum_{\substack{c>0\\ N|c}} \frac{\mathcal{S}_A(c)}{c} \quad \text{avec} \quad \mathcal{S}_A(c) := \sum_{n=1}^{+\infty} \frac{\chi(n)}{\sqrt{n}} S(1,n;c) J_1\left(\frac{4\pi\sqrt{n}}{c}\right) e^{-nx} \tag{II.23}$$

et

$$B(\chi,N)=\sum_{\substack{d>0\\(d,N)=1}}\frac{\mathcal{S}_B(d)}{d}\quad\text{avec}\quad\mathcal{S}_B(d):=\sum_{n=1}^{+\infty}\frac{\chi(n)}{\sqrt{nN}}S(1,nN^{-1};d)J_1\left(\frac{4\pi\sqrt{n}}{d\sqrt{N}}\right)e^{-nx}. \tag{II.24}$$

Plus précisément, voici de premières bornes pour $\mathcal{S}_A$ et $\mathcal{S}_B$.

Proposition II.3.1 (Bornes induites par Weil)**.** *Pour tout $N>1$ impair avec un seul diviseur premier, tout $c>0$ divisible par N et tout $d>0$ premier à N,*

$$|\mathcal{S}_A(c)|\leq\frac{2D\sqrt{N}\tau(c/N)}{\sqrt{c}},\qquad|\mathcal{S}_B(d)|\leq\frac{D\tau(d)}{\sqrt{dN}}. \tag{II.25}$$

Démonstration. Pour tout $x\in\mathbb{R}$, $|J_1(x)|\leq x/2$ d'où

$$|\mathcal{S}_A(c)|\leq\frac{2\pi}{c}\sum_{n=1}^{+\infty}|S(1,n;c)|e^{-nx},\qquad|\mathcal{S}_B(d)|\leq\frac{2\pi}{dN}\sum_{n=1}^{+\infty}|S(1,n/N;d)|e^{-nx}.$$

En utilisant les bornes de Weil, $|S(1,n;c)|\leq 2\tau(c/N)\sqrt{c}$ (par hypothèse sur N grâce à (II.12)) et $|S(1,nN^{-1};d)|\leq\tau(d)\sqrt{d}$ grâce à (II.11). Or, on a

$$\sum_{n=1}^{+\infty}e^{-nx}=\frac{1}{e^x-1}\leq\frac{1}{x}=\frac{D\sqrt{N}}{2\pi}$$

d'où les bornes pour $\mathcal{S}_A$ et $\mathcal{S}_B$. □

Ces bornes suffisent à prouver les convergences absolues des séries définissant $A(\chi,N)$ et $B(\chi,N)$, mais nous allons en trouver de meilleures pour c et d petits, basées sur une transformation d'Abel et une inégalité type Polya-Vinogradov, que nous présentons ci-dessous.

Lemme II.3.2. *Soient c,D et m trois entiers strictement positifs, et F le ppcm de c et D. Soit χ un caractère de Dirichlet quadratique de conducteur D. Alors, pour tout entier α,*

$$\left|\sum_{n=0}^{F-1}\chi(n)S(m,n;c)e^{2i\pi n\alpha/F}\right|\leq c\sqrt{D}$$

et cette somme est nulle si $(\alpha,F/(c,D))\neq 1$. Avec les mêmes notations, on a pour $c\neq D$

$$\sup_{K,K'\in\mathbb{N}}\left|\sum_{n=K}^{K'}\chi(n)S(m,n;c)\right|\leq\frac{4c\sqrt{D}}{\pi^2}(\log(Dc)+1.5).$$

Remarque II.3.1. Pour $c=1$, c'est une version de l'inégalité de Polya-Vinogradov classique pour les caractères de Dirichlet (plus grossière que le théorème 1 de [Pom11]). Pour $D=1$, c'est une inégalité analogue pour les sommes de Kloosterman qui existe probablement dans la littérature mais dont nous n'avons pas trouvé trace.

Démonstration. Soient $c'=c/(c,D)$ et $D'=D/(c,D)$. Par définition des sommes de Kloosterman,

$$\sum_{n=0}^{F-1}\chi(n)S(m,n;c)e^{2i\pi n\alpha/F}=\sum_{v\in(\mathbb{Z}/c\mathbb{Z})^*}e^{2i\pi mv^{-1}/c}\sum_{n=0}^{F-1}\chi(n)e^{2i\pi n(v/c+\alpha/F)} \tag{II.26}$$

$$=\sum_{v\in(\mathbb{Z}/c\mathbb{Z})^*}e^{2i\pi mv^{-1}/c}\left(\sum_{n'=0}^{D-1}\chi(n')\theta_v^{n'}\right)\left(\sum_{\ell=0}^{c'-1}\theta_v^{\ell D}\right) \tag{II.27}$$

avec $\theta_v := \exp(2i\pi(v/c + \alpha/F))$, car $\chi(n)$ ne dépend que de $n \bmod D$. Or, θ_v^D est une racine c'-ième de l'unité donc la somme tout à droite est nulle à moins que $\theta_v^D = 1$, c'est-à-dire si et seulement si

$$F|(D'v+\alpha)D \Longleftrightarrow c'|D'v+\alpha \Longleftrightarrow v = -(D')^{-1}\alpha \mod c'.$$

Soit I_α l'ensemble des $v \in (\mathbb{Z}/c\mathbb{Z})^*$ congrus à $(D')^{-1}\alpha$ modulo c'. Remarquons que $I_\alpha \neq \emptyset$ si et seulement si $(\alpha, c') = 1$, (en particulier $I_0 \neq \emptyset$ si et seulement si $c' = 1$), et alors

$$\sum_{n=0}^{F-1} \chi(n) S(m,n;c) e^{2i\pi n\alpha/F} = c' \sum_{v\in I_\alpha} e^{2i\pi m v^{-1}/c} \sum_{n'=0}^{D-1} \chi(n')\theta_v^{n'}.$$

La somme sur n' est une somme de Gauss associée à χ car θ_v est une racine D-ième de l'unité. Plus précisément, en définissant $G(\chi) := \sum_{n=0}^{D-1} \chi(n) e^{2in\pi/D}$, comme χ est de conducteur D, par les propriétés habituelles des sommes de Gauss, $|G(\chi)| = \sqrt{D}$ et

$$\sum_{n=0}^{F-1} \chi(n) S(m,n;c) e^{2i\pi n\alpha/F} = c' \sum_{v\in I_\alpha} e^{2i\pi m v^{-1}/c} \chi\left(\frac{D'v+\alpha}{c'}\right) G(\chi).$$

Si $(D', \alpha) \neq 1$, $\chi((D'v+\alpha)/c') = 0$ pour tout $v \in I_\alpha$ donc la somme est nulle (en particulier, elle est nulle pour $\alpha = 0$ à moins que $c = D$ car alors $D' = c' = 1$). Plus généralement, le cardinal de I_α est au plus (c, D), donc

$$\left|\sum_{n=0}^{F-1} \chi(n) S(m,n;c) e^{2i\pi n\alpha/F}\right| \leq c'(c,D)\sqrt{D} = c\sqrt{D}.$$

Passons maintenant à la seconde inégalité, suivant l'approche classique de Polya-Vinogradov. Soient $K \leq K'$ des entiers quelconques. Comme $\chi(n)S(m,n;c)$ ne dépend que de n modulo F,

$$\sum_{n=K}^{K'} \chi(n) S(m,n,c) = \frac{1}{F} \cdot \sum_{\gamma=0}^{F-1} \left[\sum_{\beta=0}^{F-1} \chi(\beta) S(m,\beta;c) e^{2i\pi\gamma\beta/F} \sum_{n=K}^{K'} e^{-2i\pi\gamma n/F}\right].$$

À droite, nous avons une somme géométrique sur n facile à borner, et la somme sur β pour $\gamma = 0$ est nulle d'après la preuve précédente car $c \neq D$. On obtient

$$\begin{aligned}\left|\sum_{n=K}^{K'} \chi(n) S(m,n,c)\right| &\leq \frac{1}{F} \cdot \sum_{\gamma=1}^{F-1} c\sqrt{D} \left|\frac{1 - e^{-2i\pi\gamma(K'-K+1)/F}}{1 - e^{-2i\pi\gamma/F}}\right| \\ &\leq \frac{c\sqrt{D}}{F} \cdot \sum_{\gamma=1}^{F-1} \frac{|\sin(\pi\gamma(K'-K+1)/F)|}{\sin(\pi\gamma/F)}.\end{aligned}$$

Il reste donc à prouver une inégalité élémentaire, à savoir que pour $F \geq 2$ et K entiers,

$$S_{K,F} := \sum_{\gamma=1}^{F-1} \frac{|\sin(\pi\gamma K/F)|}{\sin(\pi\gamma/F)} \leq \frac{4F}{\pi^2} \cdot (\log(F) + 1.5). \tag{II.28}$$

D'après le lemme 2 et la fin de la preuve du lemme 4 de [Pom11], pour tout $n \in \mathbb{N}$, $x \in \mathbb{R}$ et $F \geq 10$,

$$\sum_{j=1}^{n} \frac{\cos(jx)}{j} > -\log(2) - \frac{2}{n}, \tag{II.29}$$

$$A_{K,F} := \sum_{\gamma=1}^{F-1} \frac{1}{\sin(\pi\gamma/F)} \leq \frac{2F}{\pi} (\log(F) + 0.25). \tag{II.30}$$

En utilisant le développement en série de Fourier de $|\sin\theta|$, on a

$$\begin{aligned} S_{K,F} &= \frac{2}{\pi}\sum_{\gamma=1}^{F-1}\frac{1}{\sin(\pi\gamma/F)} - \frac{4}{\pi}\sum_{m=1}^{+\infty}\frac{1}{4m^2-1}\left(\sum_{\gamma=1}^{F-1}\frac{\cos(2\pi mK\gamma/F)}{\sin(\pi\gamma/F)}\right) \\ &= \frac{2}{\pi}A_{K,F} - \frac{4}{\pi}\sum_{m=1}^{+\infty}\frac{B_{m,K,F}}{4m^2-1}, \quad B_{m,K,F} := .\sum_{\gamma=1}^{F-1}\frac{\cos(2\pi mK\gamma/F)}{\sin(\pi\gamma/F)}. \end{aligned} \tag{II.31}$$

La borne pour $A_{K,F}$ est donnée par (II.30), il reste donc à borner $B_{m,K,F}$. Supposons que F est impair. Pour tout $x \in [0,\pi/2]$, $\sin(x) = x - \varepsilon_x x^3/6 \geq 0$ avec $\varepsilon_x \in [0,1]$, donc

$$\left|\frac{1}{\sin(x)} - \frac{1}{x}\right| \leq \frac{x}{6-x^2}.$$

On en déduit que

$$\begin{aligned} B_{m,K,F} &= 2\sum_{\gamma=1}^{(F-1)/2}\frac{\cos(2\pi mK\gamma/F)}{\sin(\pi\gamma/F)} \\ &\geq 2\sum_{\gamma=1}^{(F-1)/2}\frac{\cos(2\pi mK\gamma/F)}{\pi\gamma/F} - 2\sum_{\gamma=1}^{(F-1)/2}\frac{\pi\gamma/F}{6-(\pi\gamma/F)^2} \\ &\geq -\frac{2F}{\pi}(\log(2)+4/(F-1)) - 2\int_{1/2}^{F/2}\frac{\pi x/F}{6-(\pi x/F)^2}dx \\ &\geq -\frac{2F}{\pi}(\log(2)+4/(F-1)) - \frac{2F}{\pi}\int_0^{\pi/2}\frac{udu}{6-u^2} \\ &\geq -\frac{2F}{\pi}(0.96+4/(F-1)). \end{aligned}$$

Pour la première somme, on a utilisé (II.29) et pour la seconde, on a utilisé la convexité de la fonction $x \mapsto x/(6-x^2)$ sur l'intervalle $[0,\pi/2]$. Enfin, comme $\sum_{m=1}^{+\infty} 1/(4m^2-1) = 1/2$ (on reconnaît les termes d'une série télescopique), on déduit de (II.30) et (II.31) l'inégalité

$$S_{K,F} \leq \frac{4F}{\pi^2}(\log(F)+1.21+4/(F-1)).$$

Elle implique la majoration voulue lorsque $F \geq 15$, et on l'obtient à la main pour $F \leq 15$. Pour F pair, on procède la même manière en prenant en compte le fait que $(F-1)/2$ est remplacé par $F/2-1$ dans le calcul de $B_{m,K,F}$ ci-dessus (ce qui le simplifie légèrement). □

On déduit de ce lemme de nouvelles majorations de $|\mathcal{S}_A|$ et $|\mathcal{S}_B|$ par transformation d'Abel.

Proposition II.3.3. *Pour tous entiers positifs N, c, d avec $N|c$ et $(d,N)=1$ différents de D, on a*

$$|\mathcal{S}_A(c)| \leq 6\sqrt{D}(\log(Dc)+1.5) \quad \textit{et} \quad |\mathcal{S}_B(d)| \leq \frac{6\sqrt{D}(\log(Dd)+1.5)}{N}. \tag{II.32}$$

Démonstration. Pour tout $n \geq 0$, soit

$$A_n := \sum_{k=1}^{n}\chi(k)S(1,k;c), \qquad B_n := \sum_{k=1}^{n}\chi(k)S(1,kN^{-1};d),$$

et

$$f_A(y) = \frac{cJ_1\left(\frac{4\pi\sqrt{y}}{c}\right)}{4\pi\sqrt{y}}e^{-yx}, \qquad f_B(y) = \frac{d\sqrt{N}J_1\left(\frac{4\pi\sqrt{y}}{d\sqrt{N}}\right)}{4\pi\sqrt{y}}e^{-yx}.$$

Ainsi, par définition de $\mathcal{S}_A$ et $\mathcal{S}_B$ (formules (II.23) et (II.24)), on a

$$\mathcal{S}_A(c) = \sum_{n=1}^{+\infty}(A_n - A_{n-1})\frac{4\pi f_A(n)}{c}, \qquad \mathcal{S}_B(d) = \sum_{n=1}^{+\infty}(B_n - B_{n-1})\frac{4\pi f_B(n)}{dN}.$$

Par transformée d'Abel, on a alors

$$|\mathcal{S}_A(c)| \leq \frac{4\pi}{c}\cdot\sum_{n=1}^{+\infty}|A_n|\cdot|f_A(n) - f_A(n+1)| \quad \text{et} \quad |\mathcal{S}_B(d)| \leq \frac{4\pi}{d\sqrt{N}}\cdot\sum_{n=1}^{+\infty}|B_n|\cdot|f_B(n) - f_B(n+1)|$$

donc

$$|\mathcal{S}_A(c)| \leq \frac{16\sqrt{D}}{\pi}\operatorname{Totvar}(f_A)(\log(Dc)+1.5) \quad \text{et} \quad |\mathcal{S}_B(d)| \leq \frac{16\sqrt{D}}{\pi\sqrt{N}}\operatorname{Totvar}(f_B)(\log(Dd)+1.5)$$

grâce au lemme II.3.2, avec $\operatorname{Totvar}(f_A)$ et $\operatorname{Totvar}(f_B)$ les variations totales de f_A et f_B sur $[0,+\infty[$. D'après leur définition, ces deux variations totales sont bornées par celle de la fonction $J_1(x)/x$ on $[0,+\infty[$, qui est égale d'après la formule II.6 de [Wat22] à

$$\int_0^{+\infty}\left|\frac{J_2(x)}{x}\right|dx \leq 1.1,$$

et en arrondissant $16/\pi\cdot 1.1$ à 6, on obtient les bornes désirées. □

Lemme II.3.4. *Pour tout entier $\lambda > 0$, on a*

$$\begin{aligned}\sum_{n=1}^{\lambda}\frac{1}{n} &\leq \log(\lambda)+1,\\ \sum_{n=1}^{\lambda}\frac{\log n}{n} &\leq \frac{\log(\lambda)^2+1}{2},\\ \sum_{n=\lambda}^{+\infty}\frac{\tau(n)}{n^{3/2}} &\leq \frac{2\log(\lambda)+8}{\sqrt{\lambda}}.\end{aligned}$$

Démonstration. Les deux premières sont des applications directes de la comparaison somme-intégrale (pour la seconde, faire attention au fait que $\log(x)/x$ n'est décroissante qu'à partir de $x = e$). Pour la troisième et pour $\lambda \geq 2$, on réécrit

$$\begin{aligned}\sum_{n\geq\lambda}\frac{\tau(n)}{n^{3/2}} &= \sum_{k,\ell=1}^{+\infty}\frac{\mathbf{1}_{k\ell\geq\lambda}}{(k\ell)^{3/2}}\\ &= \sum_{k=1}^{\lceil\lambda/2\rceil-1}\sum_{\ell=\lceil\lambda/k\rceil}^{+\infty}\frac{1}{(k\ell)^{3/2}} + \sum_{k=\lceil\lambda/2\rceil}^{\lambda}\sum_{\ell=2}^{+\infty}\frac{1}{(k\ell)^{3/2}} + \sum_{k=\lambda+1}^{+\infty}\sum_{\ell=1}^{+\infty}\frac{1}{(k\ell)^{3/2}}.\end{aligned}$$

Dans la première somme du découpage, on a $\lceil \lambda/k \rceil \geq 3$, et par comparaison somme-intégrale :

$$\begin{aligned}
\sum_{k=1}^{\lceil \lambda/2 \rceil - 1} \sum_{\ell=\lceil \lambda/k \rceil}^{+\infty} \frac{1}{(k\ell)^{3/2}} &\leq \sum_{k=1}^{\lceil \lambda/2 \rceil - 1} \frac{1}{k^{3/2}} \left((\lambda/k)^{-3/2} + \int_{\lambda/k}^{+\infty} \frac{dx}{x^{3/2}} \right) \\
&\leq \sum_{k=1}^{\lceil \lambda/2 \rceil - 1} \frac{1}{k^{3/2}} \left((\lambda/k)^{-3/2} + 2(\lambda/k)^{-1/2} \right) \\
&\leq \sum_{k=1}^{\lceil \lambda/2 \rceil - 1} \lambda^{-3/2} + \frac{2}{k\sqrt{\lambda}} \\
&\leq \frac{1}{2\sqrt{\lambda}} + \frac{2}{\sqrt{\lambda}} \left(1 + \int_1^{\lambda/2} \frac{dx}{x} \right) \\
&\leq \frac{1}{2\sqrt{\lambda}} + \frac{2}{\sqrt{\lambda}} (\log(\lambda) + 1 - \log(2)) \\
&\leq \frac{1}{\sqrt{\lambda}} (2\log\lambda + 5/2 - 2\log 2)
\end{aligned}$$

Pour la seconde somme, on a

$$\begin{aligned}
\sum_{k=\lceil \lambda/2 \rceil}^{\lambda} \sum_{\ell=2}^{+\infty} \frac{1}{(k\ell)^{3/2}} &= \sum_{k=\lceil \lambda/2 \rceil}^{\lambda} \frac{\zeta(3/2) - 1}{k^{3/2}} \\
&\leq (\zeta(3/2) - 1) \left((2/\lambda)^{3/2} + \int_{\lambda/2}^{\lambda} \frac{dx}{x^{3/2}} \right) \\
&\leq \frac{\zeta(3/2) - 1}{\sqrt{\lambda}} \left(2(\sqrt{2} - 1) + \frac{2^{3/2}}{\lambda} \right).
\end{aligned}$$

Enfin, pour la troisième somme, on a

$$\begin{aligned}
\sum_{k=\lambda+1}^{+\infty} \sum_{\ell=1}^{+\infty} \frac{1}{(k\ell)^{3/2}} &= \sum_{k=\lambda+1}^{+\infty} \frac{\zeta(3/2)}{k^{3/2}} \\
&\leq \zeta(3/2) \sum_{k=\lambda+1}^{+\infty} \frac{1}{k^{3/2}} \\
&\leq \frac{2\zeta(3/2)}{\sqrt{\lambda}}.
\end{aligned}$$

On obtient donc l'inégalité

$$\begin{aligned}
\sum_{n=\lambda}^{+\infty} \frac{\tau(n)}{n^{3/2}} &\leq \frac{2\log\lambda}{\sqrt{\lambda}} + \frac{5/2 - 2\log(2) + 2(\zeta(3/2) - 1)(\sqrt{2} - 1) + 2\zeta(3/2)}{\sqrt{\lambda}} + \frac{2^{3/2}(\zeta(3/2) - 1)}{\lambda^{3/2}} \\
&\leq \frac{2\log\lambda + 7.7}{\sqrt{\lambda}} + \frac{4.6}{\lambda^{3/2}}
\end{aligned}$$

On obtient donc l'inégalité voulue lorsque $\lambda \geq 4.6/0.3$, et on vérifie par machine qu'elle reste vraie pour λ plus petit. □

On peut maintenant donner des bornes sur $|A(\chi, N)$ et $|B(\chi, N)|$.

Proposition II.3.5. *Pour $N = p$ ou p^2 avec p un nombre premier impair, on a*

$$\begin{aligned}|A(\chi,N)| &\leq \min\left(\frac{14D}{N}, \frac{\sqrt{D}}{N}\left(9\log^2(D)+6\log(D)\log(N)\right)\right)\\ |B(\chi,N)| &\leq \min\left(\frac{7D}{\sqrt{N}}, \frac{\sqrt{D}}{N}\left(9\log^2(D)+12\log(D)\log(N)+6\log^2(N)\right)+\frac{\tau(D)}{\sqrt{D}}\right)\end{aligned}$$

Démonstration. Soit $\lambda > 0$ un paramètre (non entier) et $c > 0$. On va utiliser la borne (II.32) pour $c < \lambda$ et la borne (II.25) pour $c > \lambda$, d'où

$$\begin{aligned}|A(\chi,N)| &= \left|\sum_{c>0}\frac{\mathcal{S}_A(Nc)}{Nc}\right| \leq 6\sqrt{D}\left(\sum_{c=1}^{\lfloor\lambda\rfloor}\frac{\log(DNc)+1.5}{Nc}\right)+2D\sqrt{N}\left(\sum_{c\geq\lceil\lambda\rceil}\frac{\tau(c)}{(Nc)^{3/2}}\right)\\ &\leq \frac{6\sqrt{D}}{N}\left((\log(DN)+1.5)(1+\log(\lambda))+\frac{\log(\lambda)^2+1}{2}\right)+\frac{2D}{N}\left(\frac{2\log(\lambda)+8}{\sqrt{\lambda}}\right)\end{aligned}$$

Si on choisit $\lambda < 1$, cela revient à utiliser la borne de Weil partout, et on obtient alors

$$|A(\chi,N)| \leq \frac{2D}{N}\zeta(3/2)^2 \leq \frac{14D}{N}.$$

Si $D > e^4$, on choisit $\lambda = D/e^4$ et on développe $\log(\lambda)$ dans l'inégalité ci-dessus, d'où l'inégalité

$$|A(\chi,N)| \leq \frac{\sqrt{D}}{N}(9\log(D)^2+6\log(D)\log(N)).$$

Ceci est en fait le terme dominant du développement pour le λ choisi, et on voit rapidement que les termes restants sont à somme négative. Si $D < e^4$, la majoration tient encore car elle est plus grossière que la borne précédente. On procède de même pour $B(\chi,N)$: pour $\lambda > D$, on a

$$|B(\chi,N)| \leq \frac{6\sqrt{D}}{N}\left((\log(D)+1.5)(1+\log(\lambda))+\frac{\log(\lambda)^2+1}{2}\right)+\frac{D}{\sqrt{N}}\cdot\frac{2\log(\lambda)+8}{\sqrt{\lambda}}+\frac{\tau(D)}{\sqrt{DN}}.$$

Ce dernier terme dans la somme apparaît car on n'a qu'une borne de Weil pour $d = D$. Si on choisit la borne de Weil pour $|B(\chi,N)|$, on obtient

$$|B(\chi,N)| \leq \frac{1}{\sqrt{N}}\sum_{d>0}\frac{D\tau(d)}{d^{3/2}} \leq \frac{1}{\sqrt{N}}\zeta(3/2)^2 D \leq \frac{7D}{\sqrt{N}}.$$

Pour le reste, on choisit $\lambda = DN/e^4$, et comme précédemment les termes dominants en D et N sont ceux qui donnent la borne énoncée, le reste étant négatif. □

On peut maintenant conclure sur la non-annulation des sommes pondérées.

Théorème II.1.

Soit K un corps quadratique imaginaire de discriminant $-D$ et de caractère χ. Alors, pour tout nombre premier $p > 50D^{1/4}\log D$ ne divisant pas D, $(a_1, L_\chi)_{p^2}^{+,new} \neq 0$.

Démonstration. D'après le lemme II.1.1 et la formule (II.22), on a

$$\begin{aligned}\frac{1}{4\pi}|(a_1,L_\chi)_{p^2}^{+,\mathrm{new}}| &= \frac{1}{4\pi}\left|(a_1,L_\chi)_{p^2}-\frac{(a_1,L_\chi)_p}{p-1}\right|\\ &\geq e^{-2\pi/(Dp)}-\frac{1}{p-1}-2\pi\left(|A(\chi,p^2)|+|B(\chi,p^2))|+\frac{|A(\chi,p)|}{p-1}+\frac{|B(\chi,p)|}{(p-1)}\right).\end{aligned}$$

En n'utilisant que les bornes de Weil de la proposition II.3.5, et comme pour $Dp \geq 72$, on a $e^{-2\pi/(Dp)} \geq 19/20$, on obtient

$$\begin{aligned}\frac{1}{4\pi}|(a_1, L_\chi)_{p^2}^{+,\mathrm{new}}| &> \frac{19}{20} - \frac{1}{p-1} - \frac{14\pi D}{p}\left(1 + \frac{2}{p} + \frac{2}{p-1} + \frac{1}{\sqrt{p}(1-1/p)}\right) \\ &> \frac{18}{20} - \frac{50D}{p} \quad \text{pour } p \geq 37.\end{aligned}$$

Ce terme est donc non nul dès que $p > 56D$ (on a fait des arrondis naturels dans ce calcul pour simplifier la formule finale). Cette borne, bien que pouvant être pratique dans les petits cas, n'est pas aussi bonne en fonction de D que les bornes données par la transformée d'Abel, c'est pourquoi on les donne maintenant. Avec la transformée d'Abel et les mêmes hypothèses sur p, on obtient après mise en commun des bornes et arrondis (par exemple, $\tau(D) \leq 2\sqrt{D}$ et pour nos hypothèses sur p, la contribution des termes utilisant τ est d'au plus $1/20$) :

$$\frac{1}{4\pi}\left|(a_1, L_\chi)_{p^2}^{+,\mathrm{new}}\right| \geq \frac{17}{20} - 2\pi\frac{\sqrt{D}}{p^2}\left(37\log(D)^2 + 45\log(D)\log(p) + 31\log(p)^2\right)$$

et on voit après un peu d'analyse réelle que pour $p > 50D^{1/4}\log(D)$ lorsque $D \geq 15$ le terme de droite est strictement positif, et on retrouve les petits cas avec des calculs plus fins sur les estimations précédentes. Plus précisément, pour $D = 7, 8$ ou 11, on retrouve cette borne grâce à la borne de Weil appliquée en tout terme sauf en $|B(\chi, p^2)|$ et pour $D = 3$ ou 4, on peut la retrouver grâce aux calculs de [BEN10] (qui sont en fait beaucoup plus précis pour D petit). □

Remarque II.3.2. En observant attentivement les détails des calculs, on s'aperçoit que la contribution de $|A(\chi, N)|$ est d'ordre de grandeur $N^{-1/2}$ fois plus petit que celle de $|B(\chi, N)|$ pour les bornes de Weil, alors qu'elles sont sensiblement équivalentes pour les bornes issues de la transformée d'Abel. Le phénomène parait assez difficile à expliquer, mais il permet de voir rapidement qu'on pouvait sensiblement améliorer la borne de $p \gg D$ à $p \gg D^{1/4}\log D$.

II.4 Calculs de la moyenne pondérée (cas quadratique réel)

Soit ici χ un caractère de Dirichlet quadratique pair de conducteur D, $d > 1$ et p premiers tels que $(D, dp) = 1$ et $\chi(d) = 1$. Les calculs seront pour la plus grande partie similaires à la section précédente. Les calculs détaillent (avec des constantes explicites) le théorème 1 de [LF15].

D'après le lemme II.1.1 (b), il nous faut calculer $(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d}$ et $(a_1, L_\chi)_{dp}^{\chi(p)_p,-_d}$. On fixe ici $N = dp$ ou $N = dp^2$ suivant les cas, et

$$x = \frac{2\pi}{D\sqrt{N}}.$$

D'après les équations fonctionnelles (II.2) et (II.3), pour tout $f \in S_2(\Gamma_0(N))^{-\chi(N)}$,

$$L_\chi(f) = 4\pi\int_{1/(D\sqrt{N})}^{+\infty}(f\otimes\chi)(iu)du = 2\sum_{n=1}^{+\infty}\frac{\chi(n)a_n(f)}{n}e^{-nx}.$$

On a donc, pour $\varepsilon_p := +_{p^2}$ si $N = dp^2$ (resp. $\varepsilon_p := \chi(p)_p$ si $N = dp$) :

$$(a_1, L_\chi)_N^{\varepsilon_p,-_d} = 2\sum_{n=1}^{+\infty}\frac{\chi(n)(a_1, a_n)^{\varepsilon_p,-_d}}{n}e^{-nx}.$$

On pose, pour tout diviseur Q de N tel que $(Q, N/Q) = 1$:

$$A_Q(\chi) = \sum_{n=1}^{+\infty}\frac{\chi(n)}{\sqrt{n}}e^{-nx}\sum_{\substack{c>0\\(N/Q)|c\\(c,Q)=1}}\frac{S(1, nQ^{-1}; c)}{c\sqrt{Q}}J_1\left(\frac{4\pi\sqrt{n}}{c\sqrt{Q}}\right).$$

Alors, d'après la formule des traces de Petersson (proposition II.2.2 et remarque II.2.2), on a

$$(a_1, L_\chi)_N^{\varepsilon_p,-_d} = 2\pi e^{-x} - 4\pi^2 \left(A_1(\chi) - A_{N/d}(\chi) + A_d(\chi) - A_N(\chi)\right). \tag{II.33}$$

On va chercher à montrer que chaque $A_Q(\chi)$ tend vers 0 quand p tend vers l'infini, pour obtenir la non-annulation. Tout d'abord, grâce aux bornes de Weil (II.11), on peut permuter les deux sommes dans les définitions des $A_Q(\chi)$, d'où

$$A_Q(\chi) = \sum_{\substack{c>0\\(N/Q)|c\\(c,Q)=1}} \frac{\mathcal{S}_Q(c)}{c}, \quad \mathcal{S}_Q(c) := \sum_{n=1}^{+\infty} \frac{\chi(n)}{\sqrt{Q}\sqrt{n}} e^{-nx} S(1, nQ^{-1}; c) J_1\left(\frac{4\pi\sqrt{n}}{c\sqrt{Q}}\right).$$

On en déduit les bornes suivantes :

Lemme II.4.1. *Pour tout c différent de D, on a*

$$|\mathcal{S}_Q(c)| \leq \min\left(\frac{D\sqrt{N}\tau(c)}{Q\sqrt{c}}, \frac{6\sqrt{D}}{Q}(\log(Dc) + 1.5)\right).$$

Pour $c = D$, seule la première borne est valide.

Démonstration. C'est le même raisonnement que celui des preuves des propositions II.3.1 et II.3.3 (attention aux légers changements de normalisation). □

Proposition II.4.2.
On suppose ici que $N = dp$ ou $N = dp^2$ avec d et p premiers distincts, et $p > 27$.
Pour tout Q divisant N tel que $(Q, N/Q) = 1$ et $Q < N$:

$$|A_Q(\chi)| \leq \min\left(\frac{7D\sqrt{Q}\tau(N/Q)}{N}, \frac{9\sqrt{D}}{N}\log(DN)^2\right). \tag{II.34}$$

Dans le cas où $Q = N$, il faut ajouter $\tau(D)/\sqrt{DN}$ à la seconde majoration.

Démonstration. Pour un paramètre $\lambda > 0$ non entier, on utilisera la première borne du lemme II.4.1 pour $c/(N/Q) < \lambda$, la seconde pour $c/(N/Q) > \lambda$. Dans le seul cas où $Q = N$, la seconde n'est pas utilisable pour $c = D$, et on majore alors $\mathcal{S}_N(D)/D$ par $\tau(D)/\sqrt{DN}$. A cet ajout près, on peut donc majorer pour tout Q de la manière suivante :

$$\begin{aligned}
|A_Q(\chi)| &\leq \sum_{c=1}^{\lfloor\lambda\rfloor} \frac{6\sqrt{D}/Q(\log(DNc/Q) + 1.5)}{(Nc/Q)^{3/2}} + \sum_{c=\lceil\lambda\rceil}^{+\infty} \frac{D\sqrt{N}\tau(Nc/Q)}{Q(Nc/Q)^{3/2}} \\
&\leq \frac{6\sqrt{D}}{N}\left((1.5 + \log(DN/Q))(1 + \log\lambda) + \frac{1 + \log^2\lambda}{2}\right) + \frac{D\sqrt{Q}\tau(N/Q)(2\log\lambda + 8)}{N\sqrt{\lambda}}
\end{aligned}$$

d'après le lemme II.3.4. Si on a choisi $\lambda < 1$, on n'utilise en fait que la première borne, et alors

$$|A_Q(\chi)| \leq \frac{D\sqrt{Q}\tau(N/Q)}{N}\zeta(3/2)^2 \leq \frac{7D\sqrt{Q}\tau(N/Q)}{N}.$$

Dans le cas contraire, on se donne un paramètre $\mu \geq 0$ et on fixe $\lambda = DQ/e^\mu$. En développant les termes ci-dessous (et en remplaçant $\log(DQ)$ par $\log(DN) - \log(N/Q)$), on obtient l'inégalité suivante :

$$|A_Q(\chi)| \leq \frac{\sqrt{D}}{N}\left(9\log(DN)^2 + T_1\log(DQ) + T_2\right)$$

avec

$$\begin{aligned}
T_1 &= 2e^{\mu/2}\tau(N/Q) - 12\mu - 6\log(N/Q) - 6\log N + 15, \\
T_2 &= 6(1 - \mu - \log(N/Q)/2)\log(N/Q) + 9(1 - \mu) - 6(1 - \mu)\log Q + 3(1 + \mu^2).
\end{aligned}$$

Il suffit donc de trouver un paramètre μ tel que T_1 et T_2 sont négatifs.

Pour $Q = 1$, $\tau(N) = 6$ donc $T_1 \leq 0$ équivaut à

$$e^{\mu/2} + 5/4 \leq \mu + \log N.$$

Cette inégalité est satisfaite pour $\mu = 1$ lorsque $N \geq 7$, et alors $T_2 = 6 - \log N^2/2$ est également négatif lorsque $N \geq 31$. Ceci prouve l'inégalité voulue pour $N \geq 31$ et $Q = 1$.

Pour $Q = N$, $\tau(1) = 1$ donc $T_1 \leq 0$ si

$$\frac{e^{\mu/2}}{3} + \frac{15}{6} \leq 2\mu + \log N.$$

Cette inégalité est satisfaite pour $\mu = 0$ et $N \geq 18$, et alors $T_2 \leq 0$ est automatique.

Enfin, pour $Q \neq 1, N$, $\tau(N/Q) \leq 3$ et $6\log(N/Q) \geq 6\log 2 \geq 4$, donc $T_1 \leq 0$ est impliqué par

$$e^{\mu/2} + \frac{11}{6} \leq \log N + 2\mu,$$

or cette inégalité est vraie pour $\mu = 0$ et $\mu = 1$. Or, si $\log N \geq 4$ c'est-à-dire $N \geq 55$, soit $\log Q \geq 2$, auquel cas $T_2 \leq 0$ pour $\mu = 0$, soit $\log(N/Q) \geq 2$ et alors $T_2 \leq 0$ pour $\mu = 1$, ce qui conclut la preuve.

□

Nous pouvons maintenant donner le résultat d'existence de quotients de rang zéro sur $\mathbb{Q}$ dans le cas quadratique réel.

Théorème II.2. *Soit d un nombre premier et χ un caractère quadratique pair de conducteur D. Supposons que $L_\chi = 0$ sur $S_2(\Gamma_0(d)))$.*

- *Si $D \neq 1$, alors $D \geq 5$ et*

$$(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,new} \neq 0 \quad si \quad p \geq 142 D^{1/4}\log(D).$$

  En particulier, si $\chi = 1$, $J_0(dp^2)^{+_{p^2},p-new}$ admet un quotient de rang zéro sur $\mathbb{Q}$, et si χ est le caractère de Dirichlet d'un corps quadratique réel, la tordue de $J_0(dp^2)^{+_{p^2},p-new,-_d}$ par χ admet également un quotient de rang zéro sur $\mathbb{Q}$.
- *Si $D = 1$, alors*

$$(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,new} \neq 0 \quad si \quad p \geq 125.$$

Démonstration. En majorant grossièrement $\log(dp)$ par $\log(dp^2)$ dans l'inégalité (II.34), puis en utilisant le lemme II.1.1 (b) et la formule (II.33), on a, comme $e^{-2\pi/D\sqrt{N}} - \frac{1}{p-1} \geq 9/10$ lorsque $p \geq 55$,

$$\begin{aligned}
\frac{(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,p-\mathrm{new}}}{2\pi} &\geq \frac{9}{10} - \frac{36\pi\sqrt{D}\log(Ddp^2)^2}{dp^2} - \frac{36\pi\sqrt{D}\log(Dp)^2}{dp(p-1)} \\
&\quad -2\pi\frac{\tau(D)}{\sqrt{D}dp^2} - 2\pi\frac{\tau(D)}{\sqrt{D}dp(p-1)} \\
\frac{(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,p-\mathrm{new}}}{2\pi} &\geq \frac{9}{10} - \frac{230\sqrt{D}\log(Ddp^2)^2}{dp^2} - \frac{14}{dp^2}.
\end{aligned}$$

Pour évaluer à partir de quand le terme de droite est strictement positif, on pose $C > 0$ tel que $dp^2 = C\sqrt{D}\log(D)^2$ (on suppose ici que $D \geq 2$), de sorte que

$$\begin{aligned}
\frac{\sqrt{D}\log(Ddp^2)^2}{dp^2} &= \frac{\log(D^{3/2}C\log(D)^2)^2}{C\log(D)^2} \\
&= \frac{1}{C}\left(\frac{9}{4} + \frac{3\log(C)}{\log(D)} + 6\frac{\log\log D}{\log(D)} + \frac{4\log C\log\log D}{\log(D)^2} + \frac{4\log\log(D)^2}{\log(D)^2} + (\log C)^2\right) \\
&\leq \frac{1}{C}\left(\frac{9}{4} + \frac{3\log C}{\log 5} + 1.8 + 0.8\cdot\log C + \log(C)^2\right) \\
&\leq \frac{1}{C}\left(2.7 + 3\log C + \log(C)^2\right)
\end{aligned}$$

Par analyse réelle on obtient que ceci est majoré par $8.5/2300$ lorsque $C \geq 40000$, et sous ces conditions on a bien $14/dp^2 < 0.5/10$. Comme $d \geq 2$, nous venons donc de montrer que $(a_1, L_\chi)_{dp^2}^{+_{p^2},-_d,p-\mathrm{new}} \neq 0$ lorsque $p \geq \sqrt{20000} D^{1/4} \log(D)$. On arrondit cette inégalité en $p \geq 142 D^{1/4} \log(D)$.

Dans le cas où $D = 1$, il suffit de trouver à partir de quand $230 \log(dp^2)^2/(dp^2) \leq 8.5/10$, et c'est le cas lorsque $dp^2 \geq 30000$, soit quand $p > 125$. □

Remarque II.4.1. Le cas $D = 1$ avait déja été fait par [DM97] lorsque $d \leq 2,3$ pour tout $p \geq 7$. Notre preuve étant de nature analytique, les bornes obtenues sont moins bonnes mais on obtient alors la même chose pour tout d premier et tout $p \geq 125$. Dans la veine des deux cas quadratiques (réel et imaginaire), on peut espérer prouver l'existence de quotients de rang zéro inédits sur certaines jacobiennes modulaires résistant encore à une approche « exacte ».

Les inégalités écrites ici permettraient peut-être, complétées avec des arguments semblables à la remarque de la proposition 26.15 de [IK04], de non seulement dire qu'il existe une forme modulaire f telle que $L_\chi(f) \neq 0$, mais également de minorer la proportion de telles formes modulaires parmi une base de l'espace ambiant de f (dans le cas le plus optimiste, par une constante absolue). Ceci permettrait donc de minorer la dimension des quotients d'enroulement de nos jacobiennes. La méthode existe déja pour le niveau premier p (théorème 26.3 de [IK04]), mais plusieurs difficultés supplémentaires s'ajoutent pour les niveaux p^2 et dp^2. La première est qu'elle utilise les fonctions L de carré symétrique, de définition plus complexe qu'en niveau premier, et sans expression du signe de l'équation fonctionnelle en fonction des coefficients de Fourier de f. La seconde est qu'il faut réussir encore une fois à écarter la contribution des formes anciennes, ce qui annonce des manipulations au moins aussi délicates que celles de la preuve du (b) du lemme II.1.1. Ces précisions sur les difficultés du cas avec facteur carré m'ont été aimablement données par Emmanuel Royer.

III

La méthode de Runge en toute dimension

« Horloge ! dieu sinistre, effrayant, impassible ! »
– Charles Baudelaire

Dans ce chapitre, nous présentons la méthode de Runge pour prouver des résultats de finitude du nombre de points entiers sur certaines variétés projectives. La section III.1 regroupe les définitions et résultats plus ou moins élémentaires sur les hauteurs et points entiers qui seront nécessaires pour les différentes formulations des théorèmes et leurs preuves. Comme nous l'avons déja vu dans le premier chapitre pour $X_0(p)$ (section I.7), la méthode de Runge est appliquée dans son usage classique pour des courbes algébriques, et nous y reviendrons dans la section III.2. Elle peut être rendue explicite dans une grande diversité de cas. Contrairement aux preuves qu'on peut trouver dans la littérature, nous présenterons un argument qui permet théoriquement un résultat effectif en la hauteur de départ et non pas en des hauteurs auxiliaires. Ainsi, la section III.3 montrera suivant Bilu et Parent [BP11b] comment les unités modulaires permettent un résultat effectif (et d'une précision satisfaisante) pour les courbes modulaires générales. Ensuite, nous expliquerons dans la section III.4 comment la théorie à la base de cette méthode peut se généraliser aux variétés de dimension supérieure, avec divers choix de définition de points entiers. Enfin, dans la section III.5, après de nombreux rappels sur les variétés modulaires de Siegel, qui généralisent les courbes modulaires, nous appliquerons les résultats établis dans la section III.4 à des points entiers représentant des surfaces abéliennes, avant de discuter quelques possibilités d'application futures.

III.1 Hauteurs et points entiers sur les variétés algébriques

La plupart des rappels, notations et définitions de cette section est tirée du (ou inspirée du) chapitre B de [HS00].

Soit K un corps de nombres. On note M_K (resp. M_K^∞, M_K^0) l'ensemble des places (resp. places infinies, places finies) de K.

On associe à la seule place infinie de $\mathbb{Q}$ la valeur absolue usuelle, notée $|\cdot|_\infty$, et pour tout nombre premier p, la valeur absolue $|\cdot|_p$ est définie pour tout $x \in \mathbb{Q}^*$ par

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

où $\operatorname{ord}_p(x)$ est l'unique entier tel qu'on peut écrire $x = p^{\operatorname{ord}_p(x)} a/b$ avec p ne divisant ni a ni b. Par convention, $\operatorname{ord}_p(0) = +\infty$ et $|0|_p = 0$.

Les valeurs absolues ainsi normalisées vérifient la formule du produit sur $\mathbb{Q}$, à savoir que pour tout $x \in \mathbb{Q}^*$,

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

Toute place v de K est au-dessus d'une certaine place v_0 de $\mathbb{Q}$. On appelle *degré local de* v l'entier $n_v = [K_v : \mathbb{Q}_{v_0}]$, et on normalise la valeur absolue $|\cdot|_v$ de telle sorte qu'elle prolonge $|\cdot|_{v_0}$ à K.

Pour toute place $v \in M_K$ et toute extension L de K, on a

$$\sum_{\substack{w \in M_K \\ w|v}} n_w = [L : K] n_v \tag{III.1}$$

et pour tout $x \in K^*$, on en déduit la *formule du produit*

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

Pour tout ensemble de places S de K contenant M_K^∞, on note

$$\mathcal{O}_{K,S} = \{x \in K \,|\, v(x) \geq 0 \text{ pour toute valuation } v \notin S\},$$

en particulier $\mathcal{O}_{K,M_K^\infty} = \mathcal{O}_K$.

Pour tout point $P = (x_0 : \cdots : x_n) \in \mathbb{P}^n(K)$, on définit la hauteur logarithmique de Weil de P par

$$h(P) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} n_v \log \max(|x_0|_v, \cdots, |x_n|_v) \tag{III.2}$$

Ses propriétés principales sont les suivantes :

Proposition III.1.1 (Hauteur logarithmique absolue de Weil)**.**

(*a*) *La hauteur $h(P)$ est absolue, c'est-à-dire indépendante des choix de coordonnées, et du corps de nombres K tel que $P \in \mathbb{P}^n(K)$. On peut donc la définir sur $\mathbb{P}^n(\overline{\mathbb{Q}})$.*

(*b*) *Pour tout $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ et tout $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$, $h(\sigma(P)) = h(P)$.*

(*c*) *(Propriété de Northcott) Pour toute constante C et tout entier m, l'ensemble*

$$\{P \in \mathbb{P}^n(\overline{\mathbb{Q}}) \,|\, [\mathbb{Q}(P) : \mathbb{Q}] \leq m \text{ et } h(P) \leq C\}$$

est fini.

Démonstration. L'indépendance en les choix de coordonnées est une conséquence directe de la formule du produit sur un corps de nombres. Pour le reste des propriétés, le (*a*), le (*b*) et le (*c*) sont respectivement le lemme B.2.1 (*c*) (où $H = e^{[K:\mathbb{Q}]h}$), la proposition B.2.2 et le théorème B.2.3 de [HS00]. □

On peut également définir des hauteurs sur une variété algébrique, de la manière suivante.

Définition III.1.2 (Hauteur relative à un morphisme)**.**

Soit X une variété algébrique définie sur $\overline{\mathbb{Q}}$ et $\phi : X \to \mathbb{P}^n(\overline{\mathbb{Q}})$ un morphisme. La *hauteur associée à* ϕ est la hauteur h_ϕ définie sur $X(\overline{\mathbb{Q}})$ par

$$h_\phi(P) := h(\phi(P)).$$

Dans le cas où ϕ est à fibres finies (en particulier un plongement, ou un morphisme non constant si X est une courbe), la hauteur h_ϕ a la propriété de Northcott, c'est-à-dire qu'il n'y a qu'un nombre fini de points de $X(\overline{\mathbb{Q}})$ de degré du corps de définition et de hauteur h_ϕ bornés, pour tous choix de ces bornes d'après la proposition III.1.1 (*c*).

Pour tout $x \in \overline{\mathbb{Q}}$, considérant l'injection naturelle $i : \overline{\mathbb{Q}} \subset \mathbb{P}^1(\overline{\mathbb{Q}})$, la hauteur $h(x)$ est définie pour $K = \mathbb{Q}(x)$ par

$$h(x) := h_i(x) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} n_v \log^+ |x|_v. \tag{III.3}$$

où pour tout $t \in [0, +\infty[$, $\log^+ t = \max(0, \log t)$.

Nous allons maintenant définir une topologie sur $\mathbb{P}^n(K)$ (et certains des ses ouverts notables) lorsque $(K, |\cdot|_v)$ est un corps valué.

On note $\|\cdot\|_v$ la norme sup associée $|\cdot|_v$ sur $\mathbb{A}^n(K)$ pour tout $n \in \mathbb{N}^*$.

Définition III.1.3. Soit K un corps de nombres et $n \geq 1$.

(a) Pour tout $i \in \{0, \cdots, n\}$, on note

$$U_i = \{(x_0 : \cdots : x_n) \in \mathbb{P}^n(K) \,|\, x_i \neq 0\}$$

et on appelle cet ouvert le i-ème *ouvert de coordonnées*, et φ_i la fonction de normalisation des coordonnées sur U_i, c'est-à-dire

$$\varphi_i : \left| \begin{array}{ccc} U_i & \longrightarrow & \mathbb{A}^{n+1} \\ (x_0 : \cdots : x_n) & \longmapsto & \left(\frac{x_0}{x_i}, \cdots \frac{x_n}{x_i}\right) \end{array} \right. . \tag{III.4}$$

À tout polynôme homogène $g \in \overline{K}[X_0 : \cdots : X_n]$, on associe naturellement une fonction régulière sur U_i par $g_i := g \circ \varphi_i$.

(b) Pour toute place v de M_K et tout $i \in \{0, \cdots, n\}$, on note

$$E_{i,v} := \{(P, w) \in \mathbb{P}^n(\overline{K}) \times M_{\overline{K}} \text{ tel que } w \,|\, v \text{ et } |x_i|_w = \max_{0 \leq j \leq n} |x_j|_w\}$$

qui est une partie de $U_i(\overline{K})$.

(c) Pour tout $i \in \{0, \cdots, n\}$, on note

$$E_i := \{(P, w) \in \mathbb{P}^n(\overline{K}) \times M_{\overline{K}} \,|\, |x_i|_w = \max_{0 \leq j \leq n} |x_j|_w\}.$$

Remarquons que l'union des E_i est égale à $\mathbb{P}^n(\overline{K}) \times M_{\overline{K}}$.

(d) Pour X un fermé de $\mathbb{P}^n(K)$, on utilise les mêmes notations pour les restrictions naturelles de ces parties à X.

Dans tout ce chapitre, s'il est fait mention d'un espace projectif $\mathbb{P}^n$ ambiant, les notations U_i, φ_i et E_i désignent par défaut les objets ci-dessus. On note alors d_v^i la distance sur U_i définie par

$$d_v^i(P, Q) = \|\varphi_i(P) - \varphi_i(Q)\|_v.$$

Remarquons que pour un corps de nombres K et deux points distincts P et Q de $U_i(K)$, on a $d_v^i(P, Q) = 1$ pour tout $v \in M_K$ sauf un nombre fini d'entre eux.

Définition-Proposition III.1.4 (v-topologie)**.**

Soit $(K, |\cdot|_v)$ un corps valué et $n \in \mathbb{N}^*$.

(a) Les topologies définies par les distances d_v^i sur les U_i sont compatibles sur les $U_i \cap U_j$, et il existe donc une unique topologie sur $\mathbb{P}^n(K)$, appelée la *v-topologie*, qui induit sur chaque U_i la topologie de la distance d_v^i.

(b) La v-topologie est plus fine que la topologie de Zariski sur $\mathbb{P}^n(K)$ et les fonctions rationnelles sur $\mathbb{P}^n(K)$ sont continues sur leur domaine de définition.

(c) L'espace projectif muni de la v-topologie est séparé, et même compact si K est localement compact.

Démonstration.

(*a*) Pour tous $i \neq j$, la fonction x_j/x_i est continue sur U_i donc $U_i \cap U_j$ est ouvert dans U_i (et dans U_j par rôle symétrique). De plus, sur $\mathbb{A}^{n+1}(K)$, le changement de carte induit par l'isomorphisme naturel avec $U_i \cap U_j$ est exactement l'application

$$\varphi_j \circ \varphi_i^{-1} : (x_0, \cdots, 1_i, \cdots, x_j, \cdots x_n) \mapsto \left(\frac{x_0}{x_j}, \cdots, \frac{1_i}{x_j}, \cdots, 1_j, \cdots x_n\right).$$

On voit immédiatement que celle-ci est un homéomorphisme entre les ouverts naturels de définition et d'image, ce qui prouve que les topologies induites par U_i et U_j sur $U_i \cap U_j$ sont les mêmes. L'existence d'une unique topologie sur $\mathbb{P}^n(K)$ induisant celles sur les U_i est ensuite un résultat classique de topologie.

(*b*) Les polynômes de $K[X_1, \cdots, X_n]$ sont continus sur $(\mathbb{A}^n(K), \|\cdot\|_v)$, donc les fractions rationnelles de $K(X_1, \cdots, X_n)$ le sont sur leurs domaines de définition dans $\mathbb{A}^n(K)$. On en déduit par définition des distances d_v^i qu'une fonction rationnelle ϕ sur $\mathbb{P}^n(K)$ a sa restriction continue sur chacun des ouverts de coordonnées U_i, donc est continue sur $\mathbb{P}^n(K)$. De plus, un fermé de Zariski sur $\mathbb{P}^n(K)$ est l'ensemble des zéros d'un idéal homogène, donc la réunion d'un ensemble de zéros commun d'un idéal de polynômes sur chaque U_i et il est donc également fermé pour la v-topologie sur $\mathbb{P}^n(K)$.

(*c*) Soient deux points distincts P, Q de $\mathbb{P}^n(K)$. S'ils appartiennent à un ouvert de coordonnées commun U_i, la topologie sur U_i étant métrique, on peut les séparer par deux ouverts de U_i. Sinon, prenons i et j tels que $P \in U_i \backslash U_j$ et $Q \in U_j \backslash U_i$. Par hypothèse, on peut prendre un voisinage V_P de P (resp. V_Q de Q) tel que pour tout point $(x_0 : \cdot : x_n)$ de ce voisinage, on a $x_j/x_i < 1$ (resp. $x_i/x_j < 1$), et alors $V_P \cap V_Q = \emptyset$, ce qui prouve la séparation de $\mathbb{P}^n(K)$. Ensuite, remarquons que $\mathbb{P}^n(K)$ est l'image par l'application (continue) de projection $\mathbb{A}^{n+1}(K)\backslash\{0\} \to \mathbb{P}^n(K)$ de l'union des $B \cap H_i$ où B est la boule unité de centre 0 et de rayon 1 pour la norme $\|\cdot\|_v$ et H_i l'hyperplan d'équation $x_i = 1$. Cette application de projection est continue par construction de la v-topologie, et chacun des $B \cap H_i$ est compact car fermé et borné (vu les hypothèses sur $(K, |\cdot|_v)$), donc $\mathbb{P}^n(K)$ est compact. □

Pour formaliser la plupart des résultats qui vont suivre, nous aurons besoin de la notion de M_K-constantes et des définitions associées. Ce vocabulaire est à quelques modifications près classique, et exposé par exemple dans la section B.8 de [HS00] ou la section 2.6 de [BG06].

Définition III.1.5 (M_K-constantes et parties M_K-bornées)**.**

Soit K un corps de nombres et X une variété sur K.

(*a*) Une *M_K-constante* est une famille $(c_v)_{v \in M_K}$ de réels positifs telle que $c_v = 0$ sauf pour un nombre fini de places $v \in M_K$. Les M_K-constantes forment un cône de $(\mathbb{R}^+)^{M_K}$, stable par somme et maximum sur chacune des coordonnées. Une famille $(C_v)_{v \in M_K}$ de réels est *M_K-bornée* s'il existe une M_K-constante $(c_v)_{v \in M_K}$ telle que $C_v \leq e^{c_v}$ pour tout $v \in M_K$ (en particulier, $C_v \leq 1$ pour toute place $v \in M_K$ sauf un nombre fini d'entre elles).

(*b*) Une partie E de $X(\overline{K}) \times M_{\overline{K}}$ est *subordonnée à un ouvert affine* U si $E \subset U(\overline{K}) \times M_{\overline{K}}$. Une fonction $f \in \overline{K}[U]$ est alors *M_K-bornée sur E* si la famille des $(C_v)_{v \in M_K}$ avec

$$C_v = \sup_{\substack{(P,w) \in E \\ w|v}} |f(P)|_w$$

est M_K-bornée.

(*c*) Une partie E de $X(\overline{K}) \times M_{\overline{K}}$ est *affinement M_K-bornée dans l'ouvert affine U* si toute fonction régulière $f \in \overline{K}[U]$ est M_K-bornée sur E. Il suffit de vérifier que toutes les fonctions d'une famille génératrice finie de la $\overline{K}$-algèbre $\overline{K}[U]$ sont M_K-bornées sur E.

(*d*) Une partie E de $X(\overline{K}) \times M_{\overline{K}}$ est *M_K-bornée dans X* si X admet un recouvrement fini par des ouverts affines $(U_i)_{i \in I}$ et E par des parties E_i telles que chaque E_i est affinement M_K-bornée dans l'ouvert affine U_i. Une partie E de $X(\overline{K})$ est *M_K-bornée dans X* si $E \times M_{\overline{K}}$ l'est.

Remarque III.1.1.

(*a*) Si L est une extension finie de K, on étend naturellement toute M_K-constante $(c_v)_{v\in M_K}$ à une M_L-constante renotée $(c_w)_{w\in M_L}$ par $c_w := c_v$ si w est au-dessus de v. Remarquons qu'alors, grâce à la formule (III.1)

$$\frac{1}{[L:\mathbb{Q}]}\sum_{w\in M_L} n_w c_w = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K} n_v c_v.$$

Réciproquement, si $(c_w)_{w\in M_L}$ est une M_L-constante, on peut définir $(c_v)_{v\in M_K}$ où c_v est le maximum des c_w pour w au-dessus de v. Cette fois, on a l'inégalité

$$\frac{1}{[L:\mathbb{Q}]}\sum_{w\in M_L} n_w c_w \leq \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K} n_v c_v.$$

(*b*) Pour tout $x \in K^*$, la famille $(\log^+ |x|_v)_{v\in M_K}$ est une M_K-constante. Ensuite, l'ensemble $\{(P,v) \in \mathbb{A}^n(\overline{K}) \times M_{\overline{K}} \,|\, \|P\|_v \leq 1\}$ est affinement M_K-borné dans l'ouvert affine $\mathbb{A}^n_K$. Le reste des preuves s'appuiera notablement sur ces deux exemples.

(*c*) Il est important de noter que la notion de parties M_K-bornées n'est pas stable par sous-ouvert affine : cela est dû au fait que si V est un ouvert affine de U, toute fonction régulière sur V n'est pas forcément la restriction d'une fonction régulière sur U. Cependant, il existe une certaine stabilité par recouvrement affine fini, qu'on détaille dans la proposition suivante.

Proposition III.1.6. *Soit K un corps de nombres et X une variété sur K.*

(*a*) *Si E est affinement M_K-bornée dans U, pour tout recouvrement fini de U par des ouverts affines U_i, il existe un recouvrement $(E_i)_{i\in I}$ de E tel que chaque E_i est affinement M_K-bornée dans U_i.*

(*b*) *Si E est M_K-bornée dans une variété X, pour tout recouvrement fini de X par des ouverts $(U_i)_{i\in I}$, il existe un raffinement affine $(W_\ell)_{\ell\in L}$ de ce recouvrement et un recouvrement $(E_\ell)_{\ell\in L}$ de E tel que chaque E_ℓ est affinement M_K-bornée dans W_ℓ.*

(*c*) *Si X est projective, toute partie de X est M_K-bornée dans X, et on peut donc leur appliquer le (a) et le (b).*

Démonstration. (*a*) Tout d'abord, il suffit de prouver le résultat pour un raffinement affine fini des $(U_i)_{i\in I}$. En effet, s'il est établi pour un tel raffinement $(V_{i,j})_{i,j}$ de $(U_i)_{i\in I}$ (de sorte que chaque $V_{i,j}$ est un ouvert affine de U_i), on obtient des $E_{i,j}$ affinement M_K-bornés dans les $V_{i,j}$ par hypothèse, et on pose comme E_i la réunion (finie) des $E_{i,j}$ pour tous les $V_{i,j}$ raffinant U_i. Alors, comme toute fonction régulière sur U_i l'est sur chacun des $V_{i,j}$ (en nombre fini), elle est M_K-bornée sur chaque $E_{i,j}$ donc sur leur union E_i, ainsi celle-ci est affinement M_K-bornée dans U_i.

Maintenant, une base d'ouverts affines de U est donné par les ouverts affines principaux, c'est-à-dire ceux de la forme $U_f := \{x \in U \,|\, f(x) \neq 0\}$ pour une fonction régulière $f \in K[U]$. Par quasi-compacité des ouverts affines pour la topologie de Zariski, on peut donc choisir un raffinement affine fini des $(U_i)_{i\in I}$ formé uniquement d'ouverts affines principaux, renotons-le $(U_i)_{i\in I}$ où chaque U_i est un U_{f_i}, $f_i \in K[U]$ jusqu'à la fin de la preuve. On pose ici

$$E_i = \{(P,w) \in E \mid |f_i(P)|_w = \max_j |f_j(P)|_w\}.$$

Il est clair que les E_i recouvrent E et que E_i est subordonné à U_i pour tout i car $U_i = U_{f_i}$, reste donc à montrer que E_i est affinement M_K-borné dans U_i. Posons $h_1, \cdots, h_n$ un ensemble de générateurs de $K[U]$. Alors, un ensemble de générateurs de $K[U_i]$ est $h_1, \cdots, h_n, \frac{1}{f_i}$, et comme E est affinement M_K-bornée dans U, chacune des fonctions h_j est M_K-bornée sur E, il reste donc à vérifier que $1/f_i$ l'est sur E_i. Comme l'union des U_i est U, les f_i n'ont pas de zéro commun dans $U(\overline{K})$, ainsi par le Nullstellensatz sur $K[U]$ on peut écrire

$$\sum_{i\in I} f_i g_i = 1$$

pour certains $g_i \in K[U]$. Pour toute paire $(P,w) \in E_i$, on a donc

$$1 = \left| \sum_{j \in I} g_j(P) f_j(P) \right|_w$$

d'où

$$\left| \frac{1}{f_i(P)} \right|_w = \left| \sum_{j \in I} \frac{g_j(P) f_j(P)}{f_i(P)} \right|_w \leq |I|^\delta \max_{j \in I} |g_j(P)|_w,$$

où δ vaut 1 si w est archimédienne et 0 si w est ultramétrique. Or, chacune des fonctions g_j est régulière sur U donc M_K-bornée sur E donc E_i, donc $1/f_i$ l'est, ainsi E_i est affinement M_K-borné dans U_i.

(b) Par hypothèse, il existe un recouvrement affine fini $(V_j)_{j \in J}$ de X et un recouvrement $(E_j)_{j \in J}$ de E tel que E_j est affinement M_K-borné dans V_j. Les ouverts $U_i \cap V_j$ ne sont pas forcément affines mais on en prend un raffinement affine fini $(W_{i,j,k})_{i,j,k}$ (à j fixé, les $W_{i,j,k}$ recouvrent donc V_j). D'après le (a), il existe donc un recouvrement $(E_{i,j,k})_{i,j,k}$ de E_j tel que chaque $E_{i,j,k}$ est affinement M_K-borné dans $W_{i,j,k}$ et leur réunion est E, ce qui conclut la preuve.

(c) Il suffit par inclusion de montrer que $X(\overline{K})$ est borné dans X. Choisissons un plongement projectif $X \subset \mathbb{P}^n_K$. Pour tout i, on note U_i le i-ème ouvert de coordonnées et E_i comme dans la définition III.1.3. Alors, E_i est subordonné à U_i, $K[U_i]$ est engendré par les fonctions coordonnées $x_j/x_i, j \neq i$, et pour tout $j \in \{0, \cdots, n\}$, on a

$$\sup_{w \in M_{\overline{K}}} \sup_{P \in E_i} \left| \frac{x_j}{x_i}(P) \right|_w \leq 1$$

ainsi chaque E_i est bien affinement M_K-borné dans U_i, et l'union des E_i est bien $X(\overline{K}) \times M_{\overline{K}}$. □

Remarque III.1.2. Relativement à une même variété ambiante, une partie d'un ensemble M_K-borné l'est également, en revanche la notion de partie M_K-bornée dépend de la variété ambiante. Par exemple, $U_i \subset \mathbb{P}^n$ est M_K-borné dans $\mathbb{P}^n$ mais pas dans lui-même ! C'est aussi la raison pour laquelle, dans le (b), un raffinement du recouvrement est nécessaire pour récupérer des parties affinement M_K-bornées : cela ne fonctionne pas en général avec le recouvrement de départ.

Nous allons maintenant prouver un résultat basé sur le Nullstellensatz qui sera utilisé à de nombreuses reprises par la suite.

Proposition III.1.7 (Minoration par une M_K-constante)**.** *Soit $n \geq 1$ un entier et K un corps de nombres.*

Soit X une sous-variété fermée de $\mathbb{A}^n$ et $Y_1, \cdots, Y_k$ des fermés de Zariski de X définis sur K, d'intersection Y. On pose, pour tout $i \in \{1, \cdots, k\}$, $g_{i,1}, \cdots, g_{i,j_i}$ des générateurs de l'idéal définissant Y_i dans $K[X]$ et de même $h_1, \cdots, h_s$ pour l'idéal définissant Y. Alors, il existe une M_K-constante $(c_v)_{v \in M_K}$ telle que pour tout couple $(P,w) \in X(\overline{K}) \times M_{\overline{K}}$ vérifiant $\|P\|_w \leq 1$ avec w au-dessus de v, on a la dichotomie suivante :

$$\max_{\substack{1 \leq i \leq k \\ 1 \leq j \leq j_i}} \log |g_{i,j}(P)|_w \geq -c_v \quad \textit{ou} \quad \max_{1 \leq m \leq h_s} \log |h_m(P)|_w < 0.$$

Démonstration.

(a) D'après le Nullstellensatz appliqué sur X, comme l'intersection des Y_i est Y, il existe une puissance $p \in \mathbb{N}_{>0}$ telle que chaque h_m^p appartient à l'idéal engendré par tous les $g_{i,j}$. En conséquence, il existe des polynômes $f_{i,j,m}$ de $K[X_1, \cdots, X_n]$ tels que pour tout $m \in \{1, \cdots, s\}$,

$$\sum_{\substack{1 \leq i \leq k \\ 1 \leq j \leq j_i}} g_{i,j} f_{i,j,m} = h_m^p.$$

Notons N le nombre total de générateurs $g_{i,j}$ multiplié par le nombre maximal de coefficients non nuls de chaque $f_{i,j,m}$, et pour tout $v \in M_K$, $d_v \geq 0$ le maximum des logarithmes positifs des normes $|\cdot|_v$ des coefficients des $f_{i,j,m}$. La famille $(d_v)_{v\in M_K}$ est alors une M_K-constante, et pour tout (P,w) tel que $\|P\|_w \leq 1$, on a

$$\left|\sum_{\substack{1\leq i\leq k\\ 1\leq j\leq j_i}} g_{i,j}(P)f_{i,j,m}(P)\right|_w \leq N^\delta e^{d_v} \max_{\substack{1\leq i\leq k\\ 1\leq j\leq j_i}} |g_{i,j}(P)|_w$$

où $\delta = 1$ si v est archimédienne, et 0 sinon. Alors, soit $\log|h_m(P)|_w < 0$ pour tout $m \in \{1,\cdots,s\}$, soit il existe un m tel que $\log|g_m(P)|_w \geq 0$, et alors grâce à l'inégalité précédente on a

$$1 \leq N^\delta e^{d_v} \max_{\substack{1\leq i\leq k\\ 1\leq j\leq j_i}} |g_{i,j}(P)|_w$$

donc il existe (i,j) tel que $|g_{i,j}(P)|_w \geq N^{-\delta}e^{-d_v}$. En prenant la M_K-constante $(c_v)_{v\in M_K}$ où $c_v := d_v + \delta\log N$ pour tout $v \in M_K$, on obtient le résultat voulu. □

Remarque III.1.3. (a) Intuitivement, la signification du point (b) est très naturelle : pour tout point P de $\mathbb{P}^n(\overline{K})$, soit P est w-adiquement très proche de l'intersection des Y_k (et alors la valeur de chacun des h_m normalisés en P est petite), soit non, et alors P n'est w-adiquement pas trop proche d'un des Y_i donc l'une des valeurs $|g_{i,j}\circ\varphi_m(P)|_w$ est assez grande. Le Nullstellensatz permet de quantifier cette intuition sous la forme d'un résultat de M_K-constantes.

(b) On peut remplacer le 0 dans la seconde inégalité de l'énoncé de la proposition par une M_K-constante choisie $(c_{0,v})_{v\in M_K}$, et on obtient alors une M_K-constante $(c_v)_{v\in M_K}$ dépendant de ce choix. La preuve ci-dessus fonctionne tout à fait de la même manière en remplaçant le 1 dans la dernière inégalité de la preuve par $e^{pc_{0,v}}$.

Nous allons maintenant lier les normes v-adiques de certaines fonctions avec la réduction dans un modèle entier propre. Les définitions et premières propriétés de ceux-ci ont déja été données dans la section I.2, nous allons simplement rappeler ce qu'elles sont dans le cas où $X = \mathbb{P}^n_A$ avec A un anneau de Dedekind de corps des fractions K.

Si B est un anneau de Dedekind contenant A et de corps des fractions L, pour tout point $P = (x_0 : \cdots : x_n) \in X(L)$ et tout idéal premier non nul $\mathfrak{Q}$ de B, on note $\mathfrak{P} = A \cap \mathfrak{Q}$ et après normalisation des coordonnées pour qu'elles appartiennent toutes à $B_\mathfrak{Q}$ et l'une d'elles à $B^*_\mathfrak{Q}$, on a

$$P_\mathfrak{Q} = (\overline{x_0} : \cdots : \overline{x_n}) \in \mathbb{P}^n_{k(\mathfrak{Q})},$$

qu'on peut voir comme un point de $\mathbb{P}^n(k(\mathfrak{P}))$.

La proposition suivante fait le lien définitif entre réduction des points et valuation, sa preuve est immédiate mais nous la donnons par souci de clarté.

Proposition III.1.8. *Soit K un corps de nombres et S_0 un ensemble de places de K contenant M_K^∞. Soit X un schéma projectif sur $\mathcal{O}_{K,S_0}$, considéré comme un sous-schéma fermé de $\mathbb{P}^n_{\mathcal{O}_{K,S_0}}$ pour un certain n, et Y un sous-$\mathcal{O}_{K,S_0}$-schéma fermé de X.*

Supposons que l'idéal (homogène) de $\mathcal{O}_{K,S_0}[X_0,\cdots,X_n]$ définissant Y dans $\mathbb{P}^n$ est engendré sur $\mathcal{O}_{K,S_0}$ par des polynômes homogènes $g_1,\cdots,g_s$. Soit une extension L de K et une valuation normalisée w sur L (ne prolongeant pas une valuation de S_0) associée à un certain idéal premier $\mathfrak{P}$ de L. Alors, tout point $P = (x_0 : \cdots : x_n) \in X(L)$ appartient pour un certain i à

$$E_{i,w} = \{P \in X(L) \,|\, |x_i|_w = \max_j |x_j|_w\},$$

et $P_\mathfrak{P}$ appartient à $Y_\mathfrak{p}$ (où $\mathfrak{p} = \mathfrak{P}\cap\mathcal{O}_K$) si et seulement si pour tout $j \in \{1,\cdots,s\}$,

$$|g_j\circ\varphi_i(P)|_w < 1$$

avec les notations de la définition III.1.3.

Démonstration. Pour tout idéal premier $\mathfrak{p}$ de $\mathcal{O}_{K,S_0}$, un point $\overline{P} = (\overline{x_0} : \cdots : \overline{x_n})$ appartient à $Y_{\mathfrak{p}}(\overline{k(\mathfrak{p})})$ si et seulement si ses coordonnées vérifient les équations homogènes définissant $Y_{\mathfrak{p}}$ dans $X_{\mathfrak{p}}$, mais celles-ci sont, par hypothèse, exactement les équations $g_1, \cdots, g_s$ modulo $\mathfrak{p}$.

Maintenant, pour un point $P \in X(L)$ et $\mathfrak{P}$ un idéal premier de L au-dessus de $\mathfrak{p}$ de valuation associée w, le point P appartient à un certain $E_{i,w}$ et la réduction modulo $\mathfrak{P}$ de P est alors le point

$$P_{\mathfrak{P}} = \left(\overline{\left(\frac{x_0}{x_i}\right)} : \cdots : \overline{\left(\frac{x_n}{x_i}\right)} \right) \in \mathbb{P}^n(k(\mathfrak{P})).$$

Ainsi, $P_{\mathfrak{P}}$ appartient à $Y_{\mathfrak{p}}$ si et seulement si ces coordonnées vérifient les équations $g_1, \cdots, g_s$ modulo $\mathfrak{p}$, c'est-à-dire si et seulement si pour tout $j \in \{1, \cdots, s\}$,

$$g_j \circ \varphi_i(P) \in \mathfrak{P},$$

soit si et seulement si $|g_j \circ \varphi_i(P)|_w < 1$ pour tout $j \in \{1, \cdots, s\}$. □

Remarque III.1.4. Ce résultat est intuitif, et nous conduira pour toute la suite. On peut remarquer que cette caractérisation de la réduction dans $Y_{\mathfrak{p}}$ est la même que la deuxième possibilité de la dichotomie de la proposition III.1.7, et nous pouvons donc reformuler l'intuition de celle-ci de la manière suivante : pour tout $P \in E_{i,w}$, soit P se réduit modulo $\mathfrak{P}$ dans l'intersection des sous-schémas $Y_{\mathfrak{p}}$, soit non et alors il existe un des générateurs des idéaux de définition des sous-schémas assez grands en P pour la norme $|\cdot|_w$ associée à $\mathfrak{P}$. Il est à noter que ce « assez grand » nécessite la souplesse des M_K-constantes, même si on prend des générateurs bien choisis. Cela est dû au fait que la réduction modulo $\mathfrak{p}$ de l'intersection des Y_i peut être plus petite que l'intersection des réductions modulo $\mathfrak{p}$ de chacun des Y_i : dans le cas des courbes, on peut très bien avoir deux points distincts dans $C(K)$ qui se réduisent en le même point modulo $\mathfrak{p}$, donc l'intersection n'est pas vide. Il faut donc voir l'apparition de la M_K-constante pour le « assez grand » comme la nécessité de « séparer » w-adiquement les Y_i, ce qui est plus fin que de distinguer leur réduction modulo $\mathfrak{P}$.

III.2 Principe et preuve

Le nom de la méthode de Runge est un hommage à Carl Runge, à l'origine du théorème suivant prouvé en 1887 ([Run87], p. 432).

Théorème III.1 (Runge)**.** *Si $F \in \mathbb{Z}[X,Y]$ est un polynôme irréductible de degré total n tel que sa partie homogène de degré n a au moins deux facteurs irréductibles non constants distincts dans $\mathbb{Z}[X,Y]$, alors le nombre de solutions $(x,y) \in \mathbb{Z}\times\mathbb{Q}$ de $F(x,y) = 0$ est fini et peut être explicitement borné.*

Un grand nombre de variantes et de preuves (dont certaines effectives) de ce résultat ont été données depuis lors (par exemple, [BT08] ou [HS83]). Nous nous intéresserons plutôt dans ce chapitre à la généralisation suivante, due à Bombieri [Bom83] et basée sur un résultat de Sprindzuk [Spr81].

Théorème III.2 (Bombieri-Sprindzuk)**.** *Soient C une courbe projective lisse définie sur un corps de nombres K et $\phi \in K(C)$ non constante.*

Pour toute extension finie L de K et tout ensemble de places S_L de L contenant les places archimédiennes, on dit que (L, S_L) satisfait la « condition de Runge » si $|S_L| < r_L$ où r_L est le nombre d'orbites par $\mathrm{Gal}(\overline{L}/L)$ des pôles de ϕ. Alors

(a) Si (L, S_L) satisfait la condition de Runge, l'ensemble des points « $\mathcal{O}_{L,S_L}$-entiers de (C,ϕ) » (c'est-à-dire des $P \in C(L)$ tels que $\phi(P) \in \mathcal{O}_{L,S_L}$) est fini.

(b) La réunion de tous les ensembles de points $\mathcal{O}_{L,S_L}$-entiers de (C,ϕ) où les paires (L, S_L) vérifient la condition de Runge est finie.

Dans ce théorème, le (b) implique le (a), mais le niveau d'intuition de la preuve étant différent, nous allons les prouver successivement. Ensuite, le (a) implique le théorème original de Runge : en effet, pour $F \in \mathbb{Z}[X,Y]$ irréductible, soit C la clôture projective de la courbe affine plane d'équation $F(x,y)=0$ et $\pi : C' \to C$ une normalisation de C. On définit alors $\phi = x \circ \pi$, $K = \mathbb{Q}$ et $S = \{\infty\}$. D'après la condition sur F dans le théorème original de Runge, la courbe C a deux points rationnels distincts à l'infini, donc le support de $(\phi)_\infty$ a au moins deux points rationnels distincts, ce qui permet d'appliquer le (a) du théorème III.2 et il n'existe donc qu'un nombre fini de points rationnels P de C' tels que $\phi(P) \in \mathbb{Z}$, et par surjectivité de π, cela signifie qu'il n'existe qu'un nombre fini de $(x,y) \in \mathbb{Z} \times \mathbb{Q}$ tels que $F(x,y)=0$.

La preuve originale de Bombieri ([BG06], Théorème 9.6.6) repose sur la théorie des hauteurs locales et le théorème de décomposition de Weil. Nous allons ici présenter une autre approche, initiée par Bilu dans une note non publiée et reprise dans le cas $K=\mathbb{Q}$ par Schoof ([Sch08], Chapitre 5).

La première chose à démontrer est le résultat suivant.

Proposition III.2.1. *Soit C une courbe projective lisse sur un corps de nombres K et $f_1, \cdots, f_r$ des fonctions de $\overline{K}(C)$ sans pôle commun deux à deux.*

Alors, il existe une M_K-constante $(c_v)_{v \in M_K}$ telle que pour tout $v \in M_K$ et tout (P,w) dans $C(\overline{K}) \times M_{\overline{K}}$ où w est au-dessus de v, l'inégalité

$$\log |f_i(P)|_w \leq c_v$$

est vraie pour toutes les fonctions f_i sauf peut-être une (à chaque choix de couple (P,w)).

Preuve de la proposition. Expliquons d'abord l'intuition place par place derrière ce résultat. Pour chaque $i \in \{1, \cdots, r\}$, soient $P_{i,1}, \cdots, P_{i,r_i}$ les pôles de f_i. Dans $\mathbb{P}^n(K_v)$ (compact d'après la définition-proposition III.1.4 (c)), $|f_i(P)|_v$ est grand si et seulement si P est près d'un des pôles de f_i. On choisit alors des petits voisinages autour de chacun des $P_{i,j}$ qui ne s'intersectent pas deux à deux par séparation de $\mathbb{P}^n(K_v)$, de sorte que P ne peut être près que d'une famille de pôles à la fois, donc $|f_i(P)|_v$ ne peut être grand que pour une seule valeur de i.

Voyons maintenant comment la formaliser avec la notion de M_K-constante.

Pour tout (i,j), on choisit un ouvert affine $U_{i,j}$ de $C(\overline{K})$ contenant $P_{i,j}$ mais aucun des autres pôles, et quitte à rajouter des ouverts affines supplémentaires ne contenant aucun pôle d'aucune des fonctions (qui vérifieront trivialement ce qu'on veut ci-dessous), on peut supposer que les $U_{i,j}$ recouvrent $C(\overline{K})$.

D'après la proposition III.1.6 (b) et (c), on peut alors définir des parties $E_{i,j,k}$ recouvrant $C(\overline{K}) \times M_{\overline{K}}$ telles que pour tout (i,j), la partie $E_{i,j,k}$ est affinement M_K-bornée dans un ouvert $U_{i,j,k}$ et chaque $U_{i,j,k}$ est un ouvert affine de $U_{i,j}$, leur union à (i,j) fixé recouvrant $U_{i,j}$. Alors, pour tout (i,j,k) et par construction des ouverts $U_{i,j}$, chacune des fonctions f_ℓ $(\ell \neq i)$ est régulière sur $U_{i,j}$, et donc il existe une M_K-constante $(c_v^{i,j,\ell})_v$ telle que pour tout $(P,w) \in E_{i,j,k}$,

$$\log |f_\ell(P)|_w \leq c_v^{i,j,\ell}.$$

En prenant le maximum des M_K-constantes obtenues pour chaque i et chaque $\ell \neq i$, on obtient une M_K-constante réalisant l'énoncé de la proposition. □

Preuve du théorème III.2. Prenons K, C et ϕ comme dans l'énoncé du théorème. Soit K' une extension auxiliaire finie galoisienne de K sur laquelle tous les pôles de ϕ sont définis.

(a) Soit L une extension finie de K pouvant vérifier la condition de Runge (c'est-à-dire qu'il y a au moins deux orbites de pôles de ϕ par $\mathrm{Gal}(\overline{K}/L)$) et Q, Q' deux pôles distincts de ϕ.

D'après le théorème de Riemann-Roch ([Liu02], Théorème 7.3.26 et corollaire 7.3.33), comme le degré du diviseur $2[Q]-[Q']$ est strictement positif et que celui-ci est défini sur K', il existe une fonction $g_{Q,Q'}$ de $K'(C)$ dont le seul pôle est Q et s'annulant en Q'. On note alors $f_{Q,Q',L}$ le produit de tous les conjugués par $\mathrm{Gal}(\overline{K}/L)$ de $g_{Q,Q'}$. Si on suppose que Q et Q' appartiennent à des $\mathrm{Gal}(\overline{K}/L)$-orbites distinctes (ce qu'on fera dorénavant), la fonction $f_{Q,Q',L}$ a exactement comme pôles (et de même degré) les conjugués de Q et admet comme zéros les conjugués de Q'

par $\mathrm{Gal}(\overline{K}/L)$. Pour cette version de la preuve, on a seulement besoin de savoir pour la suite qu'elle est non constante, appartient à $L(C)$ et a comme pôles les conjugués de Q, ce qu'on fait désormais. On numérote de 1 à r_L les orbites des pôles de ϕ pour la conjugaison par $\mathrm{Gal}(\overline{K}/L)$ et on renote pour tous $i \in \{1, \cdots, r_L\}$ distincts, $f_{i,L} = f_{Q_i,Q',L}$ pour un certain choix de pôle Q_i de la i-ème orbite et Q' n'appartenant pas à cette orbite.

Comme ϕ est non constante, le corps de fonctions $\overline{K}(C)$ est une extension finie de $\overline{K}(\phi)$, ainsi $g_{Q,Q'}$ est algébrique sur $\overline{K}(\phi)$. De plus, les valuations discrètes sur $\overline{K}(C)$ qui sont triviales sur $\overline{K}$ correspondent exactement aux points de $C(\overline{K})$, en associant à tout point P la valuation ord_P (voir la preuve du lemme I.6.5 de [Har77]). Ainsi, si P n'est pas un pôle de ϕ, $\mathrm{ord}_P(g_{Q,Q'}) \geq 0$ donc $g_{Q,Q'}$ appartient à l'intersection des anneaux de valuation discrète de $\overline{K}(C)$ contenant $\overline{K}[\phi]$, mais ceci est exactement la clôture intégrale de $K[\phi]$ dans $\overline{K}(C)$ d'après ([AM94], Corollaire 5.22). La fonction rationnelle $g_{Q,Q'}$ est donc entière sur $K[\phi]$ donc sur $\mathcal{O}_K[\phi]$ quitte à la multiplier par un élément non nul convenable de $\mathcal{O}_K$, ce qu'on fait désormais.

Soit P un point (L, S_L)-entier de (C, ϕ). Pour toute place v de $M_L \backslash S_L$, on a $|\phi(P)|_v \leq 1$ donc $|f_{i,L}(P)|_v \leq 1$ d'après le paragraphe précédent, il reste donc à borner les contributions des places $v \in S_L$. Pour ceci, on constate que $|\phi(P)|_v$ est grand si et seulement si P est, dans $C(L)$ muni de la v-topologie, proche d'un pôle de ϕ, ce qui à son tour signifie que $|f_{i,L}(P)|_v$ est grand pour un certain i. En restreignant suffisamment ces voisinages v-adiques (c'est-à-dire si $|\phi(P)|_v$ est vraiment grand), P ne peut être v-adiquement proche que d'une seule orbite de pôles $\mathcal{O}_i$, c'est-à-dire que $|f_{i,L}(P)|_v$ ne peut être grand que pour un seul indice i. Étant donné un point $P \in C(L)$ tel que $\phi(P) \in \mathcal{O}_{L,S_L}$, on élimine donc au plus pour chaque place v une orbite telle que $|f_{i,L}(P)|_v$ est grand, et comme (L, S_L) vérifie la condition de Runge, il reste une des fonctions $f_{i,L}$ telle que $(|f_{i,L}(P)|_v)_{v \in M_L}$ est absolument M_L-borné. La hauteur $h_{f_{i,L}}(P)$ est donc bornée, et l'ensemble des points de $C(L)$ de hauteur $h_{f_{i,L}}$ bornée est fini. On n'a qu'un nombre fini de fonctions $f_{i,L}$ ainsi construites, et chaque P appartient à un ensemble fini constitué par une des fonctions $f_{i,L}$, donc l'ensemble des points (L, S_L)-entiers de (C, ϕ) est bien fini.

(b) Pour démontrer ce résultat plus fort, on applique la proposition précédente aux fonctions $f_{1,L}, \cdots, f_{r_L,L}$ qui sont bien sans pôle commun (et définies sur L). On obtient alors une M_L-constante (qui nous donne une M_K-constante $(c_v)_{v \in M_K}$ par la remarque III.1.1 (a)) telle que pour tout $v \in M_K$ et tout prolongement de v à $\overline{K}$, pour tout point $P \in C(\overline{K})$, l'inégalité

$$\log |f_{i,L}(P)|_v \leq c_v$$

est vraie pour toutes les fonctions $f_{i,L}$ sauf peut-être une (à chaque choix de v et P). Ainsi, en éliminant encore une fois (au plus) une orbite de pôles v-proches d'un point de $C(L)$ à chaque place $v \in S_L$, il reste finalement une fonction $f_{i,L}$ telle que $|f_{i,L}(P)|_v \leq c_v$ pour toute place $v \in M_K$ en-dessous d'une place de S_L, c'est-à-dire que pour notre point de départ P tel que $\phi(P) \in \mathcal{O}_{L,S_L}$, il existe i tel que $h_{f_{i,L}}(P) \leq \sum_{v \in M_K} c_v$. De plus, les fonctions $f_{i,L}$ obtenues parcourent un ensemble fini lorsque (L, S_L) parcourt toutes les paires vérifiant la condition de Runge : ce sont des produits de conjugués de $g_{Q,Q'}$ (en nombre fini choisi au départ) or $g_{Q,Q'}$ est définie sur K' qui est une extension finie galoisienne de K. Le degré de ces extensions L est également borné par la condition de Runge car $|S_L| \geq |M_L^\infty| \geq [L : \mathbb{Q}]/2$.

Ainsi, la réunion des ensembles de l'énoncé du (b) est incluse dans une union finie d'ensembles de points de hauteur bornée (chacun pour une certaine fonction) et de degré de corps de définition borné, donc finie par la propriété de Northcott.

□

Remarque III.2.1. Dans de nombreux exemples explicites (y compris les courbes modulaires dans la section III.3), les fonctions auxiliaires utilisées peuvent être construites de manière *ad hoc* et les estimations des M_K-constantes faites par d'autres outils (par exemples des développement en série de Laurent), si bien que la méthode de Runge devient complètement effective en les hauteurs des fonctions auxiliaires de la preuve. Néanmoins, on souhaite une estimation finale complètement effective en la hauteur h_ϕ, et pour cela nous allons donner une variante de la preuve ci-dessous qui permet théoriquement de donner une borne finale sur la hauteur h_ϕ (et pour la raison ci-dessus, dans de nombreux cas concrets de manière totalement effective). Notons que la méthode utilisée

dans [Lev08] (dans le cas de la dimension 1) ou dans [Bom83] ne permettait pas de retrouver des bornes sous une forme aussi effective.

Théorème III.3 (Variante du théorème de Bombieri)**.**

Soit C une courbe projective lisse définie sur un corps de nombres K et $\phi \in K(C)$ non constante, avec les notations du théorème III.2.

Alors, on peut effectivement borner la hauteur h_ϕ de la réunion de tous les ensembles de points $\mathcal{O}_{L,S_L}$-entiers de (C,ϕ), où (L,S_L) vérifie la condition de Runge, et donc cet ensemble est fini.

Démonstration. Reprenons la preuve précédente jusqu'à la construction des $g_{Q,Q'}$ incluse. L'essentiel de la procédure repose sur le raffinement de la construction de bonnes fonctions auxiliaires.

Pour L une extension fixée de K, on définit $f_{Q,Q',L} \in L(C)$ le produit des conjugués de $g_{Q,Q'}$ par $\mathrm{Gal}(\overline{K}/L)$. Si jamais Q et Q' sont dans des orbites distinctes de pôles pour $\mathrm{Gal}(\overline{K}/L)$ (ce qu'on suppose maintenant), la fonction $f_{Q,Q',L}$ a pour uniques pôles Q et tous ses conjugués (avec même ordre) et s'annule en Q' et ses conjugués.

On numérote de 1 à r_L les orbites des pôles de ϕ par $\mathrm{Gal}(\overline{K}/L)$, et on note pour tout $i \in \{1,\cdots,r_L\}$, $f_{i,L}$ le produit de fonctions $f_{Q_i,Q'_j,L}$ où Q_i est un représentant de la i-ème orbite et Q'_j un représentant de la j-ième orbite, choisis arbitrairement. Ainsi, $f_{i,L}$ a pour seuls pôles les points de la i-ième orbite, et s'annule en chacun des autres pôles de ϕ. D'après les mêmes arguments que pour la preuve du théorème III.2, $f_{i,L}$ est entière sur $\mathcal{O}_K[\phi]$ si on a bien choisi chaque $g_{Q,Q'}$ (ce qu'on suppose désormais) et comme elle s'annule sur chaque pôle de ϕ sauf ceux de la i-ième orbite, il existe $n_i \in \mathbb{N}_{>0}$ fixé jusqu'à maintenant tel que $\phi f_{i,L}^{n_i}$ a pour seuls pôles ceux de la i-ième orbite (cette fonction pourrait même éventuellement s'annuler sur les autres pôles, mais cela est accessoire).

On peut alors appliquer la proposition III.2.1 aux fonctions $\phi f_{i,L}^{n_i}$, $i \in \{1,\cdots,r_L\}$ qui sont bien sans pôle commun (et en fait à toute famille de r_L fonctions obtenues par ce procédé, il y en a un nombre fini indépendant de (L,S_L) vérifiant la condition de Runge). On en déduit pour chacune de ces ensembles de fonctions une M_L-constante donc une M_K-constante (remarque III.1.1 (a)), et on en prend le maximum. En conséquence, il existe une M_K-constante (indépendante de (L,S_L) vérifiant la condition de Runge) $(c_v)_{v\in M_K}$ telle que pour tout $v \in M_K$ et tout $(P,w) \in C(\overline{K}) \times M_{\overline{K}}$ avec w au-dessus de v, l'inégalité

$$\log|\phi(P) f_{i,L}^{n_i}(P)|_w \leq c_v$$

est vraie sauf pour au plus un indice i (à chaque choix de v et P).

Soit donc P un point (L,S_L)-entier de (C,ϕ), cette paire vérifiant la condition de Runge. Par intégralité de $f_{i,L}$ sur $\mathcal{O}_K[\phi]$, pour chaque place $v \notin S_L$, $|f_{i,L}(P)|_v \leq 1$. Ensuite, pour chaque $v \in S_L$, il existe au plus un indice i ne vérifiant pas l'inégalité ci-dessus. Par principe des tiroirs, comme $|S_L| < r_L$, il existe un indice i tel que $\log|\phi(P) f_{i,L}^{n_i}(P)| \leq c_v$ pour toute place $v \in M_L$. Alors, par la formule du produit,

$$\begin{aligned}
0 &= \sum_{v\in M_L} n_v n_i \log|f_{i,L}(P)|_v \\
&= \sum_{\substack{v\in M_L \\ |\phi(P)|_v>1}} n_v \log|f_{i,L}^{n_i}(P)| + \sum_{\substack{v\in M_L \\ |\phi(P)|_v\leq 1}} n_v n_i \log|f_{i,L}(P)|_v \\
&\leq \sum_{\substack{v\in M_L \\ |\phi(P)|_v>1}} n_v n_i \log|f_{i,L}(P)|_v \\
&\leq \sum_{\substack{v\in M_L \\ |\phi(P)|_v>1}} n_v (c_v - \log|\phi(P)|_v) \\
&\leq \sum_{\substack{v\in M_L \\ |\phi(P)|_v>1}} n_v c_v - [L:\mathbb{Q}] h_\phi(P).
\end{aligned}$$

On obtient finalement

$$h_\phi(P) \leq \sum_{v\in M_L} \frac{n_v c_v}{[L:\mathbb{Q}]} \leq \frac{1}{[K:\mathbb{Q}]} \sum_{v\in M_K} c_v$$

d'après la remarque III.1.1 (a), d'où une borne directement sur h_ϕ comme annoncé, indépendante du couple (L, S_L) vérifiant la condition de Runge. □

Remarque III.2.2. L'idée derrière cette variante est qu'on peut construire un nombre fini de fonctions auxiliaires ϕ_i telle que pour tout point P de $C(L)$ qui est (L, S_L)-entier, il existe i tel que $|\phi_i(P)|_v$ devient mécaniquement petit dès que $|\phi(P)|_v$ devient grand, mais la formule du produit (et l'intégralité hors de S_L) interdisent justement que chaque $|\phi_i(P)|_v$ soit trop petit pour les $v \in S_L$, d'où l'interdiction pour $|\phi(P)|_v$ d'être trop grand pour toutes les places $v \in S_L$, ce qui permet de borner la hauteur : pour l'exemple des courbes modulaires, voir la proposition III.3.22 (b) et ce qui la suit.

En ce qui concerne la façon dont on construit nos fonctions à pôles disjoints, la procédure ci-dessus paraît un peu grossière, mais ne change pas qualitativement la preuve. En regardant de plus près sa nature (et l'intuition qu'on peut en avoir place par place), on remarque cependant que la situation permettant les meilleures bornes possibles est celle où les ordres n_i nécessaires sont petits, et où les $f_{i,L}$ n'ont pas des ordres de pôles trop élevés (pour améliorer la M_K-constante). Il faut donc essayer d'établir les n_i les plus petits possibles, et en un sens les constructions faites dans la section III.3 sont optimales (voir les propositions III.3.22 et III.3.18).

Enfin, on peut s'apercevoir qu'en plus d'un résultat théorique, la preuve ci-dessus fournit une méthode pour borner la hauteur des points entiers vérifiant la condition de Runge. Pour cela, il faut construire les fonctions auxiliaires, obtenir les M_K-constantes impliquées via des estimations spécifiques à la courbe pour éviter l'usage du Nullstellensatz effectif (comme on le verra dans le paragraphe III.3.2 pour les courbes modulaires), car ce dernier fournit de mauvaises bornes numériques. Enfin, on peut appliquer la formule du produit finale pour obtenir la borne. C'est ceci qu'on appelle en pratique la méthode de Runge pour les courbes.

Nous présentons finalement une version du théorème de Bombieri plus convenable pour une définition plus intrinsèque des points entiers.

Théorème III.4 (Bombieri, version modèles entiers)**.** *Soit K un corps de nombres et C une courbe projective lisse sur K.*

Supposons que C admet un modèle projectif $\mathcal{C}$ sur $\mathcal{O}_{K,S_0}$ où S_0 est un ensemble fini de places de K contenant M_K^∞.

Soit $\mathcal{P} = \{P_1, \cdots, P_r\}$ un ensemble de points distincts de $C(\overline{K})$ stable par $\mathrm{Gal}(\overline{K}/K)$. On considère, pour toute extension L de K et tout ensemble de places S_L de M_L contenant toutes les places au-dessus de S_0 :

$$(C\backslash\mathcal{P})(\mathcal{O}_{L,S_L}) = \{P \in C(L) \,|\, \forall \mathfrak{P} \notin S_L, \overline{P} \notin \overline{\mathcal{P}} \mod \mathfrak{P}\}.$$

Alors, la réunion des $(C\backslash\mathcal{P})(\mathcal{O}_{L,S_L})$ avec S_L contenant les places au-dessus de S_0 et (L, S_L) vérifiant la condition de Runge est un ensemble fini.

Démonstration. D'après le théorème de Riemann-Roch, il existe une fonction $\phi \in K(C)$ dont $\mathcal{P}$ est exactement le support des pôles. On note $\mathcal{Q}$ l'ensemble de ses zéros. Nous allons traduire la notion de points entiers de l'énoncé en termes de points entiers pour (C, ϕ) pour retrouver le théorème III.2.

On fixe un plongement projectif $\mathcal{C} \subset \mathbb{P}^n$ sur $\mathcal{O}_{K,S_0}$.

Soient $g_1, \cdots, g_s$. des générateurs homogènes de l'idéal homogène de $\mathcal{O}_{K,S_0}[X_0, \cdots, X_n]$ associé au faisceau en idéaux sur $\mathcal{O}_{K,S_0}$ défini par les prolongements des points P_i à leurs sections sur $\mathcal{O}_{K,S_0}$.

On désigne, pour tout $i \in \{0, \cdots, n\}$, par $h_{i,1}, \cdots, h_{i,j}$ des générateurs de l'idéal de $K[C \cap U_i]$ associé au diviseur de Weil des pôles de ϕ (bien défini sur K car ϕ l'est).

Pour tout $i \in \{0, \cdots, n\}$ et tout $j \in \{1, \cdots, t\}$, les fonctions $h_{i,j}/\phi$ sont régulières sur $C \cap U_i$, et elles n'ont comme zéros communs que les pôles de ϕ dans $C \cap U_i$ par construction. Ainsi, d'après la proposition III.1.7 appliquée à chaque $C \cap U_i$ (puis le maximum des M_K-constantes pris), il existe une M_K-constante $(c_v)_{v \in M_K}$ telle que pour tout $i \in \{0, \cdots, n\}$ et tout $(P, w) \in E_i$ (définition III.1.3),

$$\max_{1 \leq j \leq t} \log \left| \left(\frac{h_j \circ \varphi_i}{\phi} \right)(P) \right|_w \geq -c_v \quad \text{ou} \quad \max_{1 \leq k \leq s} \log |g_k(P)|_w < 0.$$

Maintenant, pour $P \in (C \backslash \mathcal{P})(\mathcal{O}_{L,S_L})$, le deuxième cas de la dichotomie ne peut être vérifié d'après la proposition III.1.8, donc il existe j tel que

$$\log |h_j \circ \varphi_i(P)|_w \geq \log |\phi(P)|_w - c_v.$$

Par ailleurs, sur chaque E_i, chaque fonction $h_j \circ \varphi_i$ est M_K-bornée car régulière sur U_i et car E_i est affinement M_K-borné dans U_i. En prenant le maximum $(c'_v)_{v \in M_K}$ de toutes les M_K-constantes obtenues pour chaque (i, j), on a finalement, pour tout $P \in (C \backslash \mathcal{P})(\mathcal{O}_{L,S_L})$, avec w non au-dessus de S_L :

$$\log |\phi(P)|_w \leq c_v + c'_v.$$

À part pour un nombre fini de places v, on a $c_v + c'_v = 0$ donc $\phi(P)$ est entier pour w, et pour régler les cas restants, on multiplie ϕ par une constante $\lambda \in \mathcal{O}_K$ non nulle telle que pour toute place v non archimédienne tel que $c_v + c'_v > 0$, $\log |\lambda|_v \leq -c_v - c'_v$.

Finalement, la fonction $\lambda\phi$ a pour pôles les éléments de $\mathcal{P}$, est définie sur K, et pour tout point $P \in (C \backslash \mathcal{P})(\mathcal{O}_{L,S_L})$ avec S_L contenant les places au-dessus de S_0, on a $\lambda\phi(P) \in \mathcal{O}_{L,S_L}$. On peut donc appliquer le théorème III.2 à cette fonction, d'où la finitude. □

Remarque III.2.3. Un avantage certain de la formulation du théorème précédent, en termes de points entiers, est qu'on n'a pas besoin de construire un bon modèle entier en toutes les places, ce qui peut se révéler ardu. La contrepartie est que les places non incluses dans la définition du modèle doivent automatiquement être considérées comme des « mauvaises places » au même titre que les places archimédiennes (et peuvent donc empêcher de satisfaire la condition de Runge), mais à part ceci, elles ne posent aucun problème pour l'application de la méthode de Runge.

III.3 Application aux courbes modulaires

Nous allons voir en détail dans cette section comment des fonctions propres aux courbes modulaires, les unités modulaires, permettent d'appliquer de manière systématique la méthode de Runge à ces courbes, avec la notion d'intégralité naturelle. On rappelle qu'on note ici

$\mathcal{H}$ le demi-plan de Poincaré,

$\mathcal{D}$ le domaine fondamental usuel de l'action par homographies de $\mathrm{SL}_2(\mathbb{Z})$ sur $\mathcal{H}$.

Nous allons commencer par reconstruire les fonctions de Siegel (et obtenir leurs propriétés principales), avec une approche basée sur les fonctions thêta avec caractéristique.

III.3.1 Fonctions thêta et fonctions de Siegel

Pour ce paragraphe, on pourra consulter le chapitre I de [Mum87] pour plus de contexte sur les fonctions thêta définies et leurs propriétés, notamment par rapport au groupe d'Heisenberg qu'on n'utilise qu'implicitement ici.

Définition III.3.1 (Fonctions thêta avec caractéristique)**.** Pour tous $a, b \in \mathbb{R}$, on définit la fonction holomorphe $\Theta_{a,b}$ sur $\mathbb{C} \times \mathcal{H}$ par

$$\Theta_{a,b}(z, \tau) := \sum_{n \in \mathbb{Z}} e^{i\pi(n+a)^2\tau + 2i\pi(n+a)(z+b)}.$$

La série converge absolument et uniformément sur tout compact de $\mathbb{C} \times \mathcal{H}$, et même sur tout domaine de la forme $\{(z, \tau)\ ,\ |z| \leq r, \mathrm{Im}\,\tau \geq \varepsilon\}$. La fonction $\Theta_{a,b}$ est donc holomorphe sur $\mathbb{C} \times \mathcal{H}$. En particulier, on note $\Theta := \Theta_{0,0}$.

Définition III.3.2. Pour toute fonction holomorphe f sur $\mathbb{C}\times\mathcal{H}$, on définit Sf et Tf par :

$$\begin{aligned}(Sf)(z,\tau) &:= f(z+1,\tau)\\ (Tf)(z,\tau) &:= e^{i\pi\tau+2i\pi z}f(z+\tau,\tau).\end{aligned}$$

Ainsi, S et T sont des opérateurs linéaires sur $\mathrm{Hol}(\mathbb{C}\times\mathcal{H})$.

Chacune de ces fonctions thêta est un vecteur propre pour les opérateurs S et T définis ci-dessus :

Proposition III.3.3 (Action de S et T sur les fonctions thêta)**.** *Pour tous $a,b\in\mathbb{Q}$:*
(a) Pour tous entiers $p,q\in\mathbb{Z}$,

$$\Theta_{a+p,b+q}=e^{2i\pi aq}\Theta_{a,b}.$$

(b) La fonction $\Theta_{a,b}$ vérifie $S\cdot\Theta_{a,b}=e^{2i\pi a}\Theta_{a,b}$ et $T\cdot\Theta_{a,b}=e^{-2i\pi b}\Theta_{a,b}$.
(c) Pour toute fonction holomorphe f sur $\mathbb{C}\times\mathcal{H}$ vérifiant les égalités du (b), on a $f=\lambda\Theta_{a,b}$ où pour tout $\tau\in\mathcal{H}$,

$$\lambda(\tau)=\int_0^1 f(z,\tau)e^{-i\pi a^2\tau-2i\pi a(z+b)}dz.$$

Démonstration.
(a) Soient $p,q\in\mathbb{Z}$. Alors, pour tout $(z,\tau)\in\mathbb{C}\times\mathcal{H}$,

$$\begin{aligned}\Theta_{a+p,b+q}(z,\tau) &= \sum_{n\in\mathbb{Z}} e^{i\pi(n+a+p)^2\tau+2i\pi(n+a+p)(z+q+b)}\\ &= e^{2i\pi aq}\sum_{n\in\mathbb{Z}} e^{i\pi(n+a+p)^2\tau+2i\pi(n+a+p)(z+b)}=e^{2i\pi aq}\Theta_{a,b}(z,\tau)\end{aligned}$$

après réindexation, par convergence uniforme sur tout compact.
(b) Pour tout $(z,\tau)\in\mathbb{C}\times\mathcal{H}$, on a

$$S\cdot\Theta_{a,b}(z,\tau)=\sum_{n\in\mathbb{Z}} e^{i\pi(n+a)^2\tau+2i\pi(n+a)(z+1+b)}=\Theta_{a,b+1}(z,\tau)=e^{2i\pi a}\Theta_{a,b}(z,\tau)$$

d'après le *(a)*. Ensuite,

$$\begin{aligned}T\cdot\Theta_{a,b}(z,\tau) &= e^{i\pi\tau+2i\pi z}\sum_{n\in\mathbb{Z}} e^{i\pi(n+a)^2\tau+2i\pi(n+a)(z+\tau+b)}\\ &= e^{i\pi\tau+2i\pi z}\sum_{n\in\mathbb{Z}} e^{i\pi(n+a)^2\tau+2i\pi(n+a)(z+b)+2i\pi(n+a)\tau}\\ &= e^{2i\pi z}\sum_{n\in\mathbb{Z}} e^{i\pi(n+a+1)^2\tau+2i\pi(n+a)(z+b)}\\ &= e^{-2i\pi b}\sum_{n\in\mathbb{Z}} e^{i\pi(n+a+1)^2\tau+2i\pi(n+a+1)(z+b)}\\ &= e^{-2i\pi b}\Theta_{a+1,b}(z,\tau)=e^{-2i\pi b}\Theta_{a,b}(z,\tau)\end{aligned}$$

d'après le *(a)*.
(c) Soit f une fonction holomorphe sur $\mathbb{C}\times\mathcal{H}$ ayant les mêmes valeurs propres pour S et T que $\Theta_{a,b}$. Soit $k\in\mathbb{N}_{>0}$ tel que $ka,kb\in\mathbb{Z}$. Comme la valeur propre en S est une racine k-ième de l'unité, à τ fixé, la fonction $f_\tau : z\mapsto f(z,\tau)$ est k-périodique en z et entière sur $\mathbb{C}$, elle admet donc un développement en série de Fourier

$$f_\tau(z)=\sum_{n\in\mathbb{Z}} c_n(\tau)e^{2i\pi nz/k}.$$

Par hypothèse, pour tout $n \in \mathbb{Z}$, on a

$$e^{2i\pi n/k} c_n(\tau) = e^{2i\pi a} c_n(\tau)$$

donc c_n est nul dès que n n'est pas congru à ka modulo k. De plus, on a

$$\begin{aligned} T \cdot f(z,\tau) &= e^{i\pi\tau + 2i\pi z} \sum_{n\in\mathbb{Z}} c_n(\tau) e^{2i\pi n(z+\tau)/k} \\ &= \sum_{n\in\mathbb{Z}} e^{i\pi\tau + 2i\pi n\tau/k} c_n(\tau) e^{2i\pi(n+k)z/k} \end{aligned}$$

donc pour tout $n \in \mathbb{Z}$, par identification des coefficients de Fourier,

$$e^{-2i\pi b} c_{n+k}(\tau) = e^{i\pi\tau + 2i\pi n\tau/k} c_n(\tau).$$

Ainsi, les coefficients de Fourier $c_n(\tau)$ sont uniquement déterminés par le coefficient de Fourier $c_{n_0}(\tau)$ où n_0 est le reste de la division euclidienne de ka par k. Comme ce coefficient de Fourier n'est jamais nul pour $\Theta_{a,b}$, on a bien $f = \lambda\Theta_{a,b}$ où λ est une fonction holomorphe sur $\mathcal{H}$. Reste à montrer la formule intégrale pour λ, et pour ça il suffit de prouver que l'intégrale de l'énoncé vaut 1 pour $\Theta_{a,b}$ et tout $\tau \in \mathcal{H}$, ce qu'on fait ci-dessous. On a, par convergence uniforme et permutation somme-intégrale :

$$\begin{aligned} \int_0^1 \Theta_{a,b}(z,\tau) e^{-i\pi a^2\tau - 2i\pi a(z+b)} dz &= \sum_{n\in\mathbb{Z}} \int_0^1 e^{i\pi((n+a)^2 - a^2)\tau + 2i\pi n(z+b)} dz \\ &= \sum_{n\in\mathbb{Z}} e^{i\pi((n+a)^2-a^2)\tau} \int_0^1 e^{2i\pi n(z+b)} dz \\ &= 1 \end{aligned}$$

car chaque terme pour $n \neq 0$ est nul, et le terme en $n = 0$ vaut 1. □

Ces propriétés de translation vont nous permettre d'établir une formule de modularité pour les fonctions thêta.

Définition III.3.4. Soit $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Pour tout $\tau \in \mathcal{H}$, on note

$$\gamma \cdot \tau = \frac{a\tau + b}{c\tau + d} \text{ et } j_\gamma(\tau) = c\tau + d.$$

La première formule définit une action à gauche de $\mathrm{SL}_2(\mathbb{Z})$ sur $\mathcal{H}$ et la seconde un cocycle pour cette action, c'est-à-dire que pour tous $\gamma, \gamma' \in \mathrm{SL}_2(\mathbb{Z})$,

$$j_{\gamma\gamma'}(\tau) = j_\gamma(\gamma' \cdot \tau) j_{\gamma'}(\tau).$$

Pour toute fonction holomorphe f sur $\mathcal{H}$ et toute matrice $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, la fonction f^γ est définie par

$$f^\gamma(z,\tau) := f\left(\frac{z}{j_\gamma(\tau)}, \gamma \cdot \tau\right) = f\left(\frac{z}{c\tau + d}, \frac{a\tau + b}{c\tau + d}\right),$$

ce qui définit une action à droite de $\mathrm{SL}_2(\mathbb{Z})$ sur les fonctions holomorphes sur $\mathbb{C} \times \mathcal{H}$.

On note $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $W = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Pour résoudre le conflit de notation (on a deux T),on considérera T comme la matrice ci-dessus lorsqu'elle agit à droite, et comme l'opérateur de la définition III.3.2 si T est à gauche.

Proposition III.3.5. *Pour tous $a, b \in \mathbb{Q}$,*

$$\begin{aligned}\Theta^T_{a,b}(z,\tau) &= e^{-i\pi(a^2+a)} \cdot \Theta_{a,a+b+\frac{1}{2}}(z,\tau),\\ \Theta^W_{a,b}(z,\tau) &= e^{2i\pi ab+i\pi/4} \cdot \sqrt{-\tau} \cdot e^{i\pi z^2/\tau} \cdot \Theta_{-b,a}(z,\tau)\end{aligned}$$

où dans la dernière formule, la détermination de la racine carrée est celle à valeurs dans $\mathrm{Re}(\tau) > 0$.

Remarque III.3.1. On écrit $\sqrt{-\tau}$ et non pas $\sqrt{\tau}$ pour mettre en évidence que c'est ici le cocycle $j_W(\tau)$ qui apparaît.

Démonstration. L'esprit de la démonstration est d'identifier les propriétés d'invariance des transformées de $\Theta_{a,b}$ par S et T, et d'en déduire ce qu'elles sont par la caractérisation de la proposition III.3.3 (c).

Commençons par l'image par T : par définition, $\Theta^T_{a,b}(z,\tau) = \Theta_{a,b}(z,\tau+1)$. Alors,

$$S(\Theta^T_{a,b})(z,\tau) = \Theta_{a,b}(z+1,\tau+1) = e^{2i\pi a}\Theta_{a,b}(z,\tau+1) = e^{2i\pi a}\Theta^T_{a,b}(z,\tau),$$

d'après la proposition III.3.3 (b). Ensuite,

$$\begin{aligned}T(\Theta^T_{a,b})(z,\tau) &= e^{i\pi\tau+2i\pi z}\Theta_{a,b}(z+\tau,\tau+1)\\ &= e^{i\pi\tau+2i\pi z}\Theta_{a,b}(z+(\tau+1)-1,\tau+1)\\ &= e^{i\pi\tau+2i\pi z}e^{-i\pi(\tau+1)-2i\pi(z+b)-2i\pi a}\Theta_{a,b}(z,\tau+1)\\ &= -e^{2i\pi(a+b)}\Theta^T_{a,b}(z,\tau),\end{aligned}$$

d'après la proposition III.3.3 (b). Ainsi, $\Theta^T_{a,b}$ est une fonction entière sur $\mathbb{C}\times\mathcal{H}$ qui a pour valeurs propres respectives $e^{2i\pi a}$ et $-e^{2i\pi(a+b)}$ pour S et T. D'après la proposition III.3.3 (c) appliquée à $(a, a+b+\frac{1}{2})$, elle s'écrit donc sous la forme $\Theta^T_{a,b} = \lambda(\tau)\Theta_{a,a+b+\frac{1}{2}}$, où

$$\begin{aligned}\lambda(\tau) &= \int_0^1 \Theta_{a,b}(z,\tau+1)e^{-i\pi a^2\tau-2i\pi a(z+(a+b+\frac{1}{2}))}dz\\ &= e^{i\pi a^2-2i\pi a(a+\frac{1}{2})}\int_0^1 \Theta_{a,b}(z,\tau+1)e^{-i\pi a^2(\tau+1)-2i\pi a(z+b)}dz\\ &= e^{-i\pi(a^2+a)}\end{aligned}$$

d'après la proposition III.3.3 (c) appliquée à (a,b), et donc

$$\Theta^T_{a,b} = e^{-i\pi(a^2+a)}\Theta_{a,a+b+\frac{1}{2}}.$$

Passons maintenant à l'image par W : par définition $\Theta^W_{a,b}(z,\tau) = \Theta_{a,b}(-z/\tau,-1/\tau)$. Notons $\tau' = -1/\tau$ pour faciliter l'écriture. Alors,

$$\begin{aligned}S(\Theta^W_{a,b}(z,\tau)) &= \Theta_{a,b}\left(-\frac{z+1}{\tau},\tau'\right) = \Theta_{a,b}\left(z\tau'+\tau',\tau'\right)\\ &= e^{-i\pi\tau'-2i\pi(z\tau'+b)}\Theta^W_{a,b}(z,\tau),\end{aligned}$$

d'après la proposition III.3.3 (b). Cette fois-ci, on ne tombe pas directement sur un vecteur propre pour S : définissons la fonction auxiliaire

$$\varphi_{a,b}(z,\tau) = e^{i\pi z^2\tau'}\Theta_{a,b}(z,\tau).$$

D'après le calcul précédent,

$$S(\varphi_{a,b}) = e^{-2i\pi b}\varphi_{a,b}.$$

Maintenant,

$$\begin{aligned} T(\Theta^W_{a,b})(z,\tau) &= e^{i\pi\tau+2i\pi z}\Theta_{a,b}\left(-\frac{z+\tau}{\tau},\tau'\right) = e^{i\pi\tau+2i\pi z}\Theta_{a,b}(z\tau'-1,\tau') \\ &= e^{i\pi\tau+2i\pi(z-a)}\Theta^W_{a,b}(z,\tau) \end{aligned}$$

d'après la proposition III.3.3 (b). En utilisant que $\tau\tau' = -1$ par définition, on obtient que

$$T(\varphi_{a,b}) = e^{-2i\pi a}\varphi_{a,b}.$$

D'après la proposition III.3.3 (c), on peut donc écrire

$$\varphi_{a,b}(z,\tau) = \lambda(\tau)\Theta_{-b,a}(z,\tau) \text{ avec } \lambda(\tau) = \int_0^1 \varphi_{a,b}(z,\tau)e^{-i\pi b^2\tau+2i\pi b(z+a)}dz.$$

Calculons ce facteur $\lambda(\tau)$: on a

$$\begin{aligned} \lambda(\tau) &= \int_0^1 \sum_{n\in\mathbb{Z}} e^{i\pi(n+a)^2\tau'+2i\pi(n+a)(z\tau'+b)}e^{i\pi z^2\tau'-i\pi b^2\tau+2i\pi b(z+a)}dz \\ &= \sum_{n\in\mathbb{Z}} e^{i\pi(n+a)^2\tau'+2i\pi(n+a)b-i\pi b^2\tau+2i\pi ab}\int_0^1 e^{i\pi(z^2\tau'+2(n+a)z\tau'+2bz)}dz \\ &= \sum_{n\in\mathbb{Z}} e^{i\pi(n+a)^2\tau'+2i\pi(n+a)b-i\pi b^2\tau+2i\pi ab}\int_0^1 e^{i\pi\tau'(z+(n+a-b\tau))^2-i\pi\tau'(n+a-b\tau)^2}dz \\ &= \sum_{n\in\mathbb{Z}} e^{2i\pi ab}\int_0^1 e^{i\pi\tau'(z+(n+a-b\tau))^2}dz \\ &= e^{2i\pi ab}\int_{\mathbb{R}} e^{i\pi\tau'(z-b\tau)^2}dz. \end{aligned}$$

L'intégrale ci-dessus est une gaussienne, et un peu d'analyse complexe prouve qu'en fait

$$\lambda(\tau) = e^{2i\pi ab}\int_{\mathbb{R}} e^{i\pi\tau'x^2}dx.$$

Pour finir, comme λ est une fonction holomorphe sur $\mathcal{H}$, nous allons l'évaluer pour $\tau = it, t \in \mathbb{R}$. On a

$$\lambda(it) = e^{2i\pi ab}\int_{\mathbb{R}} e^{-\pi x^2/t}dx = e^{2i\pi ab}\sqrt{t}\int_{\mathbb{R}} e^{-\pi y^2}dy = e^{2i\pi ab}\sqrt{t}.$$

On obtient donc finalement

$$\lambda(\tau) = e^{i\pi/4+2i\pi ab}\sqrt{-\tau}$$

avec la détermination de la racine annoncée, ce qui conclut la preuve. □

Pour rendre les formules de transformation plus simples, on change légèrement la normalisation des fonctions thêta de la manière suivante.

Définition III.3.6. Pour tous $a, b \in \mathbb{Q}$, définissons la fonction entière $\vartheta_{a,b}$ sur $\mathbb{C}\times\mathcal{H}$ par :

$$\vartheta_{a,b} = e^{i\pi a(1-b)}\Theta_{a-\frac{1}{2},b-\frac{1}{2}}.$$

Proposition III.3.7. *Pour tous $a, b \in \mathbb{Q}$:*

(a) Pour tous $p, q \in \mathbb{Z}$, $\vartheta_{a+p,b+q} = e^{i\pi(aq-pb+p-pq-q)}\vartheta_{a,b}$.

(b) La fonction $\vartheta_{a,b}$ est caractérisée à une fonction holomorphe sur $\mathcal{H}$ près par ses valeurs propres pour S et T, qui sont :

$$\begin{cases} S(\vartheta_{a,b}) &= -e^{2i\pi a}\vartheta_{a,b} \\ T(\vartheta_{a,b}) &= -e^{-2i\pi b}\vartheta_{a,b} \end{cases}$$

(*c*) *Les matrices T et W agissent sur les $\vartheta_{a,b}$ de la manière suivante :*

$$\begin{aligned}\vartheta^T_{a,b}(z,\tau) &= e^{i\pi/4}\cdot\vartheta_{a,a+b}\\ \vartheta^W_{a,b}(z,\tau) &= e^{3i\pi/4}\cdot\sqrt{-\tau}\cdot e^{i\pi z^2/\tau}\cdot\vartheta_{-b,a}.\end{aligned}$$

Démonstration. Les assertions (*a*), (*b*) et (*c*) se déduisent directement des propositions III.3.3 et III.3.5. Pour quand même donner un exemple de calcul, nous faisons ici le (*c*) (qui est la raison initiale de cette renormalisation) pour T et W, donnant une expression plus simple. Par définition,

$$\begin{aligned}\vartheta^T_{a,b} &= e^{i\pi a(1-b)}\Theta^T_{a-\frac12,b-\frac12} = e^{i\pi a(1-b)}e^{-i\pi((a-\frac12)^2+a-\frac12)}\Theta_{a-\frac12,a+b-\frac12}\\ &= e^{i\pi/4}e^{-i\pi(a^2+ab-a)}e^{-i\pi a(1-a-b)}\vartheta_{a,a+b}\\ &= e^{i\pi/4}\vartheta_{a,a+b}.\end{aligned}$$

De la même manière,

$$\begin{aligned}\vartheta^W_{a,b}(z,\tau) &= e^{i\pi a(1-b)}e^{2i\pi(a-\frac12)(b-\frac12)}e^{i\pi/4}\sqrt{-\tau}e^{\frac{i\pi z^2}{\tau+1}}\Theta_{a+b-\frac12,a-\frac12}(z,\tau)\\ &= e^{3i\pi/4}e^{i\pi(a-ab+2ab-a-b)}\sqrt{-\tau}e^{\frac{i\pi z^2}{\tau}}e^{i\pi b(1-a)}\vartheta_{-b,a}(z,\tau)\\ &= e^{3i\pi/4}\sqrt{-\tau}e^{\frac{i\pi z^2}{\tau}}\vartheta_{-b,a}(z,\tau).\end{aligned}$$

□

Quelque chose d'important est à remarquer dans le (*c*) : les vecteurs ligne $(a,a+b)$ et $(-b,a)$ sont respectivement les produits matriciels $(a,b)T$ et $(a,b)W$. Ceci nous permet de donner dans la proposition suivante la formule de transformation complète des fonctions thêta avec caractéristique.

Théorème III.5. *Pour tous $a',b'\in\mathbb{Q}$ et toute matrice $\gamma=\begin{pmatrix}a&b\\c&d\end{pmatrix}\in\mathrm{SL}_2(\mathbb{Z})$,*

$$\vartheta^\gamma_{a',b'}(z,\tau)=\zeta_8(\gamma)\sqrt{c\tau+d}\,e^{\frac{i\pi cz^2}{c\tau+d}}\vartheta_{(a',b')\cdot\gamma}(z,\tau),$$

avec ζ_8 une racine huitième de l'unité dépendant de γ et du choix de la racine de $c\tau+d$ (donc définie à ±1 près). Son carré, noté ζ_4, est un morphisme de groupes de $\mathrm{SL}_2(\mathbb{Z})$ dans μ_4 qui vaut i en W et T.

Remarque III.3.2. Ce résultat est une version plus précise de la proposition III.5.22 (*b*) en dimension supérieure (avec un choix de normalisation légèrement différent).

Démonstration. Cette formule est vraie pour $\gamma=T,W$ et leurs inverses d'après la proposition III.3.7 (*e*). Or, ces matrices engendrent $\mathrm{SL}_2(\mathbb{Z})$, et il suffit donc de montrer que cette formule de transformation est stable par produit. On reconnaît que $c\tau+d=j_\gamma(\tau)$, et les propriétés de cocycle de j_γ permettent facilement de voir que le seul terme à observer de près est le terme exponentiel. En fait, si on définit pour $\gamma=\begin{pmatrix}a&b\\c&d\end{pmatrix}\in\mathrm{SL}_2(\mathbb{Z})$, $c_\gamma=c$ par commodité, on doit démontrer, pour la fonction

$$\psi_\gamma(z,\tau)=e^{i\pi\frac{c_\gamma z^2}{j_\gamma(\tau)}}$$

la propriété de cocycle

$$\psi_{\gamma\gamma'}(z,\tau)=\psi_\gamma(\gamma'(z,\tau))\psi_{\gamma'}(z,\tau)$$

ce qui revient à prouver que

$$\frac{c_{\gamma\gamma'}}{j_{\gamma\gamma'}(\tau)}=\frac{c_\gamma}{j_\gamma(\gamma'\tau)j_{\gamma'}(\tau)^2}+\frac{c_{\gamma'}}{j_{\gamma'}(\tau)}.$$

Cette dernière égalité se vérifie après passage au même dénominateur en utilisant la propriété de cocycle et $\det(\gamma')=1$.

Pour ce qui est de la racine huitième de l'unité, le défaut de définition vient du fait que $j_\gamma(\tau)$ peut avoir n'importe quel argument, et il est donc impossible d'avoir une détermination unique de la racine dans tous les cas. Une situation éclairante est celle de W puisque $W^4 = 1$ alors que le facteur est $e^{-i\pi/4}$. Par contre ζ_4 est bien un morphisme de groupes car il n'y a aucune ambiguïté quand aux autres termes, la question de détermination de la racine ne se posant plus. □

Pour retrouver les fonctions de Siegel, nous avons besoin des développements en produit des fonctions thêta, que nous allons établir ci-dessous. Tout d'abord, trouvons les zéros de $\Theta_{a,b}(z,\tau)$ pour tous $a, b \in \mathbb{Q}$.

Lemme III.3.8.
Pour tout $\tau \in \mathcal{H}$,
(a) Les zéros (tous simples) de $\Theta_{a,b}(\cdot,\tau)$ sont les points de la forme $(a+p+1/2)\tau+(b+q+1/2)$ avec $p, q \in \mathbb{Z}$.
(b) Les zéros (tous simples) de $\vartheta_{a,b}(\cdot,\tau)$ sont les points de la forme $(a+p)\tau+(b+q)$ avec $p, q \in \mathbb{Z}$.

Démonstration. Le (b) est un corollaire immédiat du (a), qu'on va montrer ci-dessous.
Un fait non utilisé avant mais évident par manipulation de somme est que pour tous $a, b \in \mathbb{Q}$,

$$\Theta_{a,b}(z,\tau) = e^{i\pi a^2\tau+2i\pi a(z+b)}\Theta(z+a\tau+b,\tau),$$

donc il suffit de connaître les zéros de Θ pour obtenir le résultat. Par invariance par S et T, l'ensemble des zéros de $\Theta_\tau := \Theta(\cdot,\tau)$ est clairement $(\mathbb{Z}+\mathbb{Z}\tau)$-périodique, et nous allons montrer que le seul zéro de Θ_τ dans le domaine fondamental $D_\tau = \{x+y\tau, (x,y) \in [0,1]^2\}$ est $(1+\tau)/2$, ce qui conclura la preuve. Soit $\mathcal{P}$ un parallélogramme orienté $\mathcal{P} = z_0 + [0, 1, 1+\tau, \tau, 0]$ ne rencontrant aucun zéro de Θ. Par le théorème des résidus, si n est le nombre de zéros de Θ dans le domaine fondamental de bord $\mathcal{P}$,

$$n = \frac{1}{2i\pi}\int_{\mathcal{P}} \frac{f'(z)}{f(z)}dz.$$

Les intégrales sur les chemins $[z_0+\ell, z_0+\ell+\ell\tau]$ et $[z_0+\ell, z_0]$ se compensent car Θ_τ est 1-périodique. Ensuite, comme Θ_τ est τ-périodique,

$$\frac{\Theta'_\tau(z+\tau)}{\Theta_\tau(z+\tau)} = -2i\pi z + \frac{\Theta'_\tau(z)}{\Theta_\tau(z)}.$$

En conséquence,

$$n = \frac{1}{2i\pi}\int_0^1 2i\pi dt = 1.$$

Reste maintenant à trouver cet unique zéro, mais on remarque que $\Theta_{1/2,1/2}$ est impaire à τ fixé donc s'annule en 0, et donc Θ_τ s'annule en $(1+\tau)/2$. □

Remarque III.3.3. Cette preuve analytique a son pendant algébrique, qu'on peut trouver dans la proposition III.5.24.

Nous venons de trouver les zéros des $\Theta_{a,b}$. Un développement en produit de $\Theta_{a,b}$ se doit donc de faire apparaître ces zéros. Remarquons que pour $m \in \mathbb{Z}$,

$$e^{i\pi(2m+1)\tau-2i\pi z} = -1 \iff (2m+1)\tau - 2z = -2n-1$$

pour un certain $n \in \mathbb{Z}$, ce qui équivaut à $z-(1+\tau)/2 = m\tau+n$ avec m, n entiers. On s'attend donc à ce qu'un développement en produit de Θ fasse apparaître des termes de la forme du terme de gauche. Le résultat précis est le suivant :

Proposition III.3.9. *La fonction Θ admet le développement en produit suivant :*

$$\Theta(z,\tau) = \prod_{m=1}^{+\infty} \left(1 - e^{2i\pi m\tau}\right)\left(1 + e^{i\pi(2m-1)\tau + 2i\pi z}\right)\left(1 + e^{i\pi(2m-1)\tau - 2i\pi z}\right)$$

Ceci, avec les notations $q = e^{i\pi\tau}$ et $w = e^{i\pi z}$, équivaut à l'identité fameuse appelée triple produit de Jacobi :

$$\sum_{n\in\mathbb{Z}} q^{n^2} w^{2n} = \prod_{m=1}^{+\infty} \left(1 - q^{2m}\right)\left(1 + q^{2m-1}w^2\right)\left(1 + q^{2m-1}w^{-2}\right)$$

Démonstration. C'est le théorème 352 p. 372 de [HW08], où on trouvera également une preuve élémentaire de cette identité de Jacobi. □

Cette formule de développement en produit amène des développements pour les fonctions $\Theta_{a,b}$ et $\vartheta_{a,b}$, que nous écrivons ici.

Proposition III.3.10. *Pour tous $a, b \in \mathbb{Q}$, avec les notations $q = e^{i\pi\tau}$ et $w = e^{i\pi z}$:*

$$\Theta_{a,b}(z,\tau) = e^{2i\pi ab} q^{a^2} w^{2a} \prod_{m=1}^{+\infty} \left(1 - q^{2m}\right)\left(1 + e^{2i\pi b} q^{2(m+a)-1} w^2\right)\left(1 + e^{-2i\pi b} q^{2(m-a)-1} w^{-2}\right).$$

En particulier, pour tous $a, b \in \mathbb{Q}$ non tous les deux entiers,

$$\vartheta_{a,b}(0,\tau) = i e^{i\pi(a-1)b} q^{(a-\frac{1}{2})^2} \prod_{m=1}^{+\infty} \left(1 - q^{2m}\right)\left(1 - e^{2i\pi b} q^{2(m+a-1)}\right)\left(1 - e^{-2i\pi b} q^{2(m-a)}\right).$$

Nous pouvons maintenant retrouver les unités modulaires de Siegel grâce à ces fonctions.

Définition III.3.11 (Fonction êta de Dedekind)**.**
La fonction η de Dedekind est, avec la notation $q_\tau = e^{i\pi\tau}$, la fonction holomorphe sur $\mathcal{H}$ définie par le produit convergent

$$\eta(\tau) = q_\tau^{1/12} \prod_{m=1}^{+\infty} (1 - q_\tau^{2m}).$$

Elle vérifie, pour tout $\tau \in \mathcal{H}$,

$$\eta(\tau + 1) = e^{\frac{i\pi}{12}} \eta(\tau) \quad \text{et} \quad \eta(-1/\tau) = \sqrt{-i\tau}\,\eta(\tau)$$

([Apo90], Théorème 3.1). On remarque immédiatement que grâce à la proposition III.3.10, on a le produit

$$\eta(\tau)^3 = (\vartheta_{0,\frac{1}{2}} \vartheta_{\frac{1}{2},0} \vartheta_{\frac{1}{2},\frac{1}{2}})(0,\tau).$$

Proposition III.3.12 (Unités modulaires de Siegel)**.** *Pour tout $(a,b) \in \mathbb{Q}^2 \backslash \mathbb{Z}^2$, avec les notations de ([KL81], section 2.1) pour les fonctions de Siegel $g_{a,b}$, on a*

$$g_{a,b} = -i \frac{\vartheta_{a,b}(0,\tau)}{\eta(\tau)}.$$

En conséquence, $g_{a,b} = g_{-a,-b}$, et pour tout $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

$$(g_{a,b}^6)^\gamma = \zeta_4(\gamma)^2 g_{(a,b)\gamma}^6$$

et si $N(a,b) \in \mathbb{Z}^2$, $g_{a,b}^{2N}$ ne dépend que de $(a,b) \mod \mathbb{Z}$ et $g_{a,b}^{12N}$. Enfin, $g_{a,b}$ a le développement en produit

$$g_{a,b}(\tau) = e^{i\pi(a-1)b} q^{B_2(a)} \prod_{m=1}^{+\infty} \left(1 - e^{2i\pi b} q^{2(m+a-1)}\right)\left(1 - e^{-2i\pi b} q^{2(m-a)}\right)$$

où $B_2(X) = X^2 - X + 1/6$ est le deuxième polynôme de Bernoulli.

Démonstration. La première et la dernière égalité s'observent avec les développements en produit (attention à la convention de [KL81] où $q = e^{2i\pi\tau}$ contrairement à celle qu'on utilise dans ce paragraphe). Elle entraîne directement que pour tout $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

$$(g_{a,b}^6)^\gamma = -\frac{(\vartheta_{a,b}^6)^\gamma(0,\tau)}{(\eta^6)^\gamma(\tau)} = -\frac{\zeta_4^3(\gamma)j_\gamma(\tau)^6\vartheta_{(a,b)\gamma}^6(0,\tau)}{\zeta_4(\gamma)j_\gamma(\tau)^6\eta^6(\tau)} = \zeta_4(\gamma)^2 g_{(a,b)\gamma}^6.$$

Pour le reste, ce sont des conséquences immédiates de la proposition III.3.7. □

Définition III.3.13. L'ordre rationnel d'une fonction $g : \mathcal{H} \to \mathbb{C}$ en $i\infty$, s'il existe, est l'unique rationnel ℓ tel que $\lim_{\tau\to i\infty} e^{-2i\pi\ell\tau} g(\tau)$ existe et est non nulle.

La proposition III.3.12 implique immédiatement le corollaire suivant.

Corollaire III.3.1. *Soit $\underline{a} \in \mathbb{Q}^2 \setminus \mathbb{Z}^2$. La fonction $g_{\underline{a}}$ n'a ni pôle ni zéro sur $\mathcal{H}$, et son ordre rationnel en la pointe ∞ vaut*

$$\mathrm{ord}_\infty g_{\underline{a}} = \frac{B_2(\{a_1\})}{2}$$

où $\{.\}$ désigne la partie fractionnaire d'un réel.

On peut au passage montrer un résultat agréable, à savoir que les unités modulaires engendrent avec le j-invariant les corps de fonctions des courbes modulaires.

Proposition III.3.14. *Soit $N \geq 1$ un entier fixé. Alors, le corps de fonctions de $X(N)$ sur $\mathbb{C}$ n'est autre que*

$$\mathbb{C}(X(N)) = \mathbb{C}(j, \{g_{a,b}^{12N}, (a,b) \in D_N\}) = \mathbb{C}(j, g_{1/N,0}^{12N}, g_{0,1/N}^{12N}).$$

où D_N est l'ensemble des éléments d'ordre exactement N de $\mathbb{Q}^2/\mathbb{Z}^2$, quotienté par $\{\pm 1\}$. On a également

$$\mathbb{C}(X_1(N)) = \mathbb{C}(j, g_{0,1/N}^{12N}).$$

Démonstration. On sait déja que $\mathbb{C}(X(1)) = \mathbb{C}(j)$. Le groupe $\mathrm{SL}_2(\mathbb{Z})$ agit sur $\mathbb{C}(X(N))$ par son action à droite naturelle, car $\Gamma(N)$ est distingué dans $\mathrm{SL}_2(\mathbb{Z})$. Il agit également naturellement sur D_N, et le noyau de son action est $\Gamma(N)$. Le morphisme de groupes

$$\varphi : \left| \begin{array}{ccc} \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\pm I & \longrightarrow & \mathrm{Aut}(\mathbb{C}(X(N))) \\ \gamma & \longmapsto & (f \mapsto f \circ \gamma) \end{array} \right.$$

est donc bien défini et injectif, car son action permute les $g_{a,b}^{12N}$, $(a,b) \in D_N$ par la proposition III.3.12, et ceux-ci sont distincts : on peut choisir des représentants (a,b) de D_N dans $[0,1/2]^2$, et alors le produit infini est de la forme

$$\left(1 - e^{2i\pi b}q^{2a}\right)^{12N}\left(1 - e^{-2i\pi b}q^{2(1-a)}\right)^{12N}\prod_{m=2}^{+\infty}(1 + O(q^{2m-1}))^{24N}$$

donc le développement limité de $g_{a,b}^{12N}$ est de la forme :

$$g_{a,b}^{12N}(\tau) = e^{12i\pi(a-1)b}q^{12N(a^2-a+1/6)}\left(1 - 12Ne^{2i\pi b}q^{2a} - 12Ne^{-2i\pi b}q^{2(1-a)} + 0(q^{3a})\right).$$

Alors, si $0 \leq a < \frac{1}{2}$, le terme dominant de droite est $(1 - 12Ne^{2i\pi b}q^{2a})$, ce qui caractérise bien a et b dans $[0, \frac{1}{2}]^2$, et si $a = \frac{1}{2}$, on peut se servir de la puissance de q en facteur pour éliminer les autres possibilités. Ainsi, les $g_{a,b}^{12N}$ sont distincts (et même non proportionnels deux à deux), ce qui prouve l'injectivité du morphisme de groupes. Le groupe des automorphismes de corps de $\mathbb{C}(X(N))$ contient donc $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\pm I$ et par construction, les éléments du corps invariants par ce

sous-groupe constituent le corps $\mathbb{C}(X(1)) = \mathbb{C}(j)$: on en déduit que l'extension $\mathbb{C}(X(N))/\mathbb{C}(X(1))$ est galoisienne de groupe de Galois $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\pm I$ par le lemme d'Artin.

Ensuite, dans $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\pm I$, les éléments fixant $(0, 1/N)$ sont exactement les éléments de $\pm\Gamma_1(N)$, ceux fixant à la fois $(0, 1/N)$ et $(1/N, 0)$ sont $\pm\,\mathrm{Id}$, et donc

$$\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(j, g_{1/N,0}^{12N}, g_{0,1/N}^{12N})) = \{1\} \quad \text{et} \quad \mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(j, g_{0,1/N}^{12N})) = \varphi(\Gamma_1(N))$$

Or, par définition de φ, $\mathbb{C}(X_1(N)) = \mathbb{C}(X(N))^{\varphi(\Gamma_1(N))}$. Par correspondance de Galois, on a donc

$$\mathbb{C}(X(N)) = \mathbb{C}(j, g_{1/N,0}^{12N}, g_{0,1/N}^{12N}) \quad \text{et} \quad \mathbb{C}(X_1(N)) = \mathbb{C}(j, g_{0,1/N}^{12N}).$$

□

Enfin, démontrons les propriétés d'intégralité des $g_{\underline{a}}$, également fondamentales pour la suite. Commençons par une proposition utile.

Proposition III.3.15 ([KL81], Chapitre 2.2, Lemme 2.1)**.** *Soit $f : \mathcal{H} \to \mathbb{C}$ une fonction $\Gamma(N)$-invariante holomorphe sur $\mathcal{H}$, telle que pour tout $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, le développement de Fourier de $f_{|\alpha}$ est une série algébrique entière en $q^{1/N}$. Alors, f est entière sur $\mathbb{Z}[j]$.*

Démonstration. Tout d'abord, comme f est invariante par $\Gamma(N)$, l'ensemble S des $f_{|\alpha}$ où α parcourt $\mathrm{SL}_2(\mathbb{Z})$ est fini. Considérons alors le polynôme

$$P(X) := \prod_{g \in S} (X - g) = \sum_{n \in \mathbb{N}} a_n X^n.$$

Ses coefficients sont des fonctions holomorphes $a_n : \mathcal{H} \to \mathbb{C}$ qui sont $\mathrm{SL}_2(\mathbb{Z})$-invariantes et sans pôle sur $\mathcal{H}$. En effet, f est sans pôle sur $\mathcal{H}$ donc les éléments de S aussi, et les a_n sont des fonctions symétriques élémentaires en les éléments de S qui sont permutés par $\mathrm{SL}_2(\mathbb{Z})$, donc ils sont invariants par $\mathrm{SL}_2(\mathbb{Z})$. En conséquence, ce sont tous des éléments de $\mathbb{C}[j]$. Mais comme les développements de Fourier des éléments de S sont à coefficients algébriques entiers par hypothèses, ceux des a_n aussi. On en déduit donc que $a_n \in \overline{\mathbb{Z}}[j]$ pour tout $n \in \mathbb{N}$, c'est-à-dire que $P(X)$ est à coefficients dans $\overline{\mathbb{Z}}[j]$. La fonction f est annulée par ce polynôme unitaire, donc entière sur $\mathbb{Z}[j]$. □

Proposition III.3.16. *Soit $N \geq 2$ un entier. Soit $\underline{a} = (a_1, a_2) \in (N^{-1}\mathbb{Z})^2 \setminus \mathbb{Z}^2$ et ζ_N une racine N-ième de l'unité. Les fonctions $g_{\underline{a}}$ et $(1 - \zeta_N)g_{\underline{a}}^{-1}$ sont entières sur $\mathbb{Z}[j]$.*

Démonstration. Nous allons utiliser le critère précédent. La fonction $g_{\underline{a}}^{12N}$ est $\Gamma(N)$-modulaire et holomorphe sur $\mathcal{H}$, et son développement en série de Fourier en ∞ est algébrique entier (admettant un développement en produits à coefficients entiers algébriques). Il en est de même des conjugués de $g_{\underline{a}}^{12N}$ car ceux-ci sont de la forme $g_{\underline{a}.\alpha}^{12N}$ et $\underline{a}.\alpha$ vérifie les mêmes hypothèses que $\underline{a}$. Le critère assure alors que $g_{\underline{a}}^{12N}$ est entière sur $\mathbb{Z}[j]$, donc $g_{\underline{a}}$ aussi.

Pour la suite, regardons de plus près le développement de $g_{\underline{a}}$. On peut se ramener au cas où $0 \leq a_1 < 1$ car à racine de l'unité près, les $g_{\underline{a}}$ sont égaux quand on translate $\underline{a}$ par un couple d'entiers. Dans la proposition III.3.12, les deux facteurs devant le produit infini sont eux-mêmes inversibles dans $\mathbb{Z}((q^{1/N^2}))$, il reste donc à regarder le produit infini. Tous les coefficients sont des entiers algébriques, la série sera donc inversible dans $\overline{\mathbb{Z}}$ si le premier coefficient l'est. Or, ce premier coefficient, qu'on trouve dans le facteur $n = 0$, vaut $(1 - \exp(2i\pi a_2))$ si $a_1 = 0$ et 1 sinon. Dans le second cas, la série de $g_{\underline{a}}^{-1}$ est à coefficients entiers algébriques, dans le premier c'est celle de $(g_{\underline{a}}/(1 - \exp(2i\pi a_2)))^{-1}$, soit celle de $(1 - \zeta_N)g_{\underline{a}}^{-1}$ (les différents choix $(1 - \zeta_N)$ sont égaux à un inversible de $\mathbb{Z}[\zeta_N]$ près). On ne peut donc pas appliquer le critère à $g_{\underline{a}}^{-12N}$ mais seulement à $(1 - \zeta_N)^{12N} g_{\underline{a}}^{-12N}$ qui est bien à coefficients algébriques entiers quel que soit $\underline{a}$ de dénominateur au plus N. Le résultat en découle. □

Remarque III.3.4. Grâce à cette proposition, nous bénéficierons ici de circonstances légèrement plus favorables pour les fonctions auxiliaires dans le théorème de Runge : en effet, l'intégralité à une constante près connue sur $\mathbb{Z}[j]$ permet de borner (dans les deux sens) $|g_{a,b}(P)|_v$ dès que $j(P)$ est v-entier. Il est à noter que ceci est particulier aux pointes de courbes modulaires : le théorème de Riemann-Roch ne permet pas de garantir que les zéros et les pôles d'une fonction non constante soient contenus dans un ensemble fini prescrit.

En résumé, les fonctions de Siegel ont donc les propriétés suivantes (qu'on peut retrouver dans la section 2.1 de [KL81]).

Proposition III.3.17.
Soient $N \geq 1$. Les fonctions de Siegel $g_{a,b} : \mathcal{H} \to \mathbb{C}$ vérifient pour tout $(a,b) \in (N^{-1}\mathbb{Z})^2 \setminus \mathbb{Z}^2$:

- $g_{-a,-b} = -g_{a,b}$.
- *La fonction $g_{a,b}^{12N}$ est invariante par le groupe de congruences $\Gamma(N)$ et ne dépend que de la classe de (a,b) modulo $\mathbb{Z}^2$.*
- *Pour tout $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ vu comme une homographie de $\mathcal{H}$, $g_{a,b}^{12N} \circ \gamma = g_{(a,b)\gamma}^{12N}$.*
- *Les fonctions $g_{a,b}$ et $(1-\zeta_N)g_{a,b}^{-1}$ sont entières sur l'anneau $\mathbb{Z}[j]$ avec ζ_N une racine N-ième primitive de l'unité.*
- *Pour tout $\tau \in \mathcal{H}$ et tout rationnel ℓ, avec la notation $q^\ell = e^{2i\pi\ell\tau}$:*

$$g_{a,b}(\tau) = -q^{\frac{B_2(a)}{2}} e^{i\pi b(a-1)} \prod_{n=0}^{+\infty} (1 - e^{2i\pi b} q^{n+a})(1 - e^{-2i\pi b} q^{n+1-a}),$$

avec $B_2(X) = X^2 - X + 1/6$ le second polynôme de Bernoulli.
- *L'ordre rationnel de $g_{a,b}$ en $i\infty$ est $B_2(\{a\})/2$ avec $\{a\}$ la partie fractionnaire de a, et elle n'a ni pôles ni zéros sur $\mathcal{H}$.*

Fixons dorénavant l'entier $N \geq 1$.

Pour tout $(a,b) \in (N^{-1}\mathbb{Z})^2 \setminus \mathbb{Z}^2$, on note $u_{a,b}$ la fonction sur $X(N)$ induite par $g_{a,b}^{12N}$, ceci pour bien signifier la différence entre l'étude sur la courbe modulaire et l'étude sur le demi-plan de Poincaré.

La proposition précédente et la proposition 1.3 du chapitre 2 de [KL81] donnent le résultat suivant.

Proposition III.3.18. *Pour tout $(a,b) \in (N^{-1}\mathbb{Z})^2 \setminus \mathbb{Z}^2$:*

(a) $u_{a,b} = u_{-a,-b}$ et ne dépend que de $(a,b) \mod \mathbb{Z}^2$.

(b) Les fonctions $u_{a,b}$ sont $\mathbb{Q}(\zeta_N)$-rationnelles sur $X(N)$, et on peut choisir un isomorphisme $G = \mathrm{Gal}(\mathbb{Q}(\zeta_N)(X(N))/\mathbb{Q}(j)) \cong \mathrm{GL}_2\left(\mathbb{Z}/N\mathbb{Z}\right)/\pm 1$ tel que pour tout $\sigma \in G$, $u_{(a,b)}^\sigma = u_{(a,b).\sigma}$.

(c) Les fonctions $u_{a,b}$ et $(1-\zeta_N)^{12N} u_{a,b}^{-1}$ sont entières sur $\mathbb{Z}[j]$.

(d) Les fonctions $u_{a,b}$ n'admettent de pôles et de zéros qu'en les pointes de $\mathbb{Z}[j]$.

Remarque III.3.5. Du point de vue de la section III.2, on voudrait faire jouer aux fonctions $u_{a,b}$ le rôle des fonctions $f_{Q,Q'}$ pour la méthode de Runge explicite : elles ont les pôles bien contenus dans les pôles de j (c'est-à-dire les pointes), sont définies sur le corps de définition de $X(N)$ et entières sur $\mathbb{Z}[j]$. Le problème est que telles quelles, ces fonctions ne « sélectionnent » pas correctement leurs zéros et leurs pôles, mais nous verrons dans la proposition III.3.22 qu'on peut fabriquer les $f_{Q,Q'}$ à partir de produits de puissances des $u_{a,b}$ pour justement compenser ce défaut de sélection précise.

III.3.2 Estimations analytiques des unités modulaires près des pointes et Runge pour les courbes modulaires

On note ici, pour tout $\tau \in \mathcal{H}$ et tout $\ell \in \mathbb{Q}$, $q_\tau^\ell := e^{2i\pi\tau/\ell}$ (attention, cela diffère de certaines conventions de la section précédente).

L'invariant modulaire $j : \mathcal{H} \to \mathbb{C}$ est défini par $j(\tau) := (12c_2(\tau))^3/\Delta(\tau)$ avec c_2 et Δ qui admettent les q-développements en $i\infty$

$$c_2(\tau) = \frac{(2\pi)^4}{12}\left(1 + 240\sum_{n=1}^{\infty}\frac{n^3q^n}{1-q^n}\right) \quad \text{et} \quad \Delta(\tau) = (2\pi)^{12}q\prod_{n=1}^{\infty}(1-q^n)^{24}.$$

Pour $N \geq 1$ un entier, on cherche à borner la hauteur de certains points entiers de $(X(N), j)$.

Dans cette sous-section, nous allons citer (en général sans les redémontrer) les estimations analytiques de [BP11b], dans le but de rendre explicite les M_K-constantes prédites par la preuve de la méthode de Runge. Pour cela, les unités modulaires construites dans la section précédentes seront le bon outil.

Commençons par la partie la plus naturelle, à savoir les estimations en les places archimédiennes (il suffit de le faire pour la valeur absolue usuelle). La proposition suivante et son corollaire ([BP11b], Proposition 2.1 et corollaire 2.2), établis grâce au q-développement de j, nous fournissent des estimations analytiques du j-invariant en fonction de $q_\tau = e^{2i\pi\tau}$.

Proposition III.3.19 ([BP11b], Proposition 2.1 et corollaire 2.2)**.**
Pour tout $\tau \in \mathcal{D} + \mathbb{Z}$:
(a) On a $|\log|q_\tau|| \leq \log(|j(\tau)| + 2400)$.
(b) On a $|j(\tau)| \leq 3500$ à moins que $|q_\tau| < 0.001$.
(c) Si $|j(\tau)| > 3500$, alors $|j(\tau) - q^{-1}| \leq 1100$ et $3/2|j(\tau)| \geq |q_\tau^{-1}| \geq 1/2|j(\tau)|$.

Nous allons maintenant également estimer les unités modulaires près des pointes. Pour une pointe c de $X(N)$ et $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ tel que $\gamma\cdot\infty = c$, on définit le paramètre q_c sur $\mathcal{H}$ par $q_c = q\circ\gamma^{-1}$.

Proposition III.3.20 ([BP11b], Proposition 2.3 et corollaire 2.4)**.**
Soit $N \in \mathbb{N}^*$, $(a,b) \in \mathbb{Q}^2\backslash\mathbb{Z}^2$ tel que $N(a,b) \in \mathbb{Z}^2$ et c une pointe de $X(N)$. Alors,
(a) Pour tout $\tau \in \mathcal{D} + \mathbb{Z}$,

$$|\log|g_{a,b}| - \ell_{a,b}\log|q_\tau|| \leq \log N,$$

où $\ell_{a,b}$ est l'ordre d'annulation de $g_{a,b}$ en la pointe ∞.
(b) Pour tout $\tau \in \mathcal{H}$,

$$|\log|g_{a,b}(\tau)|| \leq \frac{1}{12}(\log|j(\tau)| + 2400) + \log N.$$

(c) Pour tout $\tau \in \mathcal{H}$, soit $|j(\tau)| \leq 3500$, soit $|q_c(\tau)| < 0.001$ pour une certaine pointe c de $X(N)$ et alors, pour tout $(a,b) \in N^{-1}\mathbb{Z}^2\backslash\mathbb{Z}^2$,

$$|\log|g_{a,b}| + \ell_{a,b,c}\log|j(\tau)|| \leq \log N + \frac{\log 3}{12},$$

où $\ell_{a,b,c}$ est l'ordre d'annulation de $g_{a,b}$ en la pointe c.

Esquisse de preuve. Pour le (a), on utilise le développement en produit en q_τ pour établir le résultat, pour le (b), quitte à remplacer τ par $\gamma\tau$ et (a,b) par $(a,b)\gamma$, grâce à la formule de transformation, il suffit de montrer ceci pour tout (a,b) et tout $\tau \in \mathcal{D}$. C'est alors une conséquence du (a), de la proposition III.3.19 (a) et du fait que $|\ell_{a,b}[\leq 1/12$. Pour le (c), pour tout $\tau \in \mathcal{H}$, il existe γ tel que $\tau' = \gamma\cdot\tau \in \mathcal{D}$, et alors si $|j(\tau')| > 3500$, on a $|q(\tau')| < 0.001$ c'est-à-dire que $|q_c(\tau)| < 0.001$ pour $c = \gamma^{-1}\cdot\infty$. Pour le reste, on réutilise le (a), ainsi que le (c) de la proposition III.3.19 appliqué à τ'. □

Cherchons maintenant à évaluer $|j(P)|_v$ et les $|g_{a,b}(P)|_v$ près des pointes, lorsque v est non-archimédienne. Fixons v une place non-archimédienne pour le corps de nombres K contenant $\mathbb{Q}(\zeta_N)$. On se place sur $X(N)(\overline{K_v})$. Pour toute pointe c de $X(N)$, d'après le corollaire 2.5 du chapitre VII de [DR73], la complétion de $X(N)$ sur $\mathbb{Z}[\zeta_N]$ le long de la section c est isomorphe à

$\mathrm{Spec}(\mathbb{Z}[\zeta_N][[q_c^{1/N}]])$. En conséquence, les q-développements complexes du j-invariant et des unités modulaires restent valides sur un voisinage v-analytique de chaque pointe, et un point P de $X(N)(\overline{K_v})$ se réduit en la pointe c si et seulement si le paramètre q_c est bien défini en P et que $|q_c(P)|_v < 1$. On définit donc $\Omega_{c,v}$ comme l'ensemble des points $P \in X(N)(\overline{K_v})$ se réduisant en c modulo v. Ces voisinages v-adiques joueront le rôle des voisinages v-adiques utilisés dans la preuve du théorème III.2.

On en déduit le résultat d'estimation analytique non-archimédienne suivant (c'est la proposition 2.5 de [BP11b]).

Proposition III.3.21.

Pour une place non-archimédienne v de K, et une pointe c de $X(N)$, notons $\Omega_{c,v}$ l'ensemble des points de $X(N)(\overline{K_v})$ se réduisant en c modulo v. Alors :

(a) Pour tout $P \in X(N)(\overline{K_v})$, $|j(P)|_v > 1$ si et seulement s'il existe c tel que $P \in \Omega_{c,v}$ et alors $|q_c(P)|_v < 1$.

(b) Pour tout $P \in \Omega_{c,v}$, $|j(P)|_v = |q_c(P)^{-1}|_v$.

(c) Pour tout $(a,b) \in N^{-1}\mathbb{Z}^2\backslash\mathbb{Z}^2$, toute pointe c et tout $P \in \Omega_{c,v}$,

$$|\log|u_{a,b}(P)|_v - 12N\ell_{a,b,c}\log|q_c(P)|_v| = 0 \text{ si } v(N) = 0,$$

et si $v|p|N$ avec p premier, on a

$$|\log|u_{a,b}(P)|_v - 12N\ell_{a,b,c}\log|q_c(P)|_v| \leq \frac{12N\log p}{p-1}.$$

Les mêmes inégalités sont valables avec $|\log|u_{a,b}(P)|_v + 12N\ell_{a,b,c}\log|j(P)|_v|$.

Démonstration. Par une propriété classique du j-invariant, $|j(P)|_v > 1$ si et seulement si P se réduit en une pointe c modulo v, d'où le (a) par construction de $\Omega_{c,v}$. Pour le (b), on utilise le q_c-développement de $g_{a,b}$, qui est convergent car $|q_c(P)|_v < 1$, et on observe immédiatement que chacun des termes du produit infini est de valeur absolue 1, sauf si $v|p|N$, auquel cas seul le terme $n = 0, a = 0$ ne l'est pas, et on a

$$1 \geq |1 - e^{2i\pi b}|_v \geq p^{-1/(p-1)}.$$

□

Remarque III.3.6. Ici, notre choix de voisinage v-adiques $\Omega_{c,v}$, une propriété de j et la modularité des courbes permettent d'obtenir le résultat du (a) sous une forme très simple : en général, on devrait avoir l'exponentielle d'une M_K-constante d'après la proposition III.2.1. C'est donc un des cas où le passage par un modèle entier et l'existence de bonnes fonctions permet d'alléger la théorie utilisée.

Vu la proposition précédente, on définit la M_K-constante $(c_v)_{v\in M_K}$ par $c_v = 0$ si $v(N) = 0$ et v est non-archimédienne, $c_v = 12N\log N + N\log 3$ si v est archimédienne, et $c_v = 12N(\log p)/(p-1)$ si $v|p|N$.

Sans entrer dans les détails, nous allons maintenant donner l'idée de la construction générale des unités modulaires voulues pour toute courbe modulaire. Soit G un sous-groupe de $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, et $X_G = X(N)/(G \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}))$ la courbe modulaire associée. Pour K un corps contenant le corps de définition de X_G, le groupe $G' = \mathrm{Gal}(K(\zeta_N)(X(N))/K(X_G))$ est un sous-groupe de $G/\pm 1$ qui contient $(G \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}))/\pm 1$. Alors, pour tout $(a,b) \in N^{-1}\mathbb{Z}^2\backslash\mathbb{Z}^2$, on note (en se rappelant la proposition III.3.18)

$$w_{a,b} = \prod_{\sigma\in G'} u_{(a,b)\sigma} = \prod_{\sigma\in G'} u_{(a,b)}^{\sigma} \in K(X_G).$$

Il est clair que ces fonctions héritent des propriétés des unités modulaires $u_{a,b}$ (en particulier l'intégralité et l'absence de zéros ou de pôles hors des pointes). La base théorique pour cette construction est que si $\mathcal{C}(G,K)$ est l'ensemble des orbites de pointes de X_G pour l'action de $\mathrm{Gal}(\overline{K}/K)$,

le groupe engendré par les diviseurs de $w_{a,b}$ est de rang maximal, c'est-à-dire $|\mathcal{C}(G,K)|-1$: c'est une conséquence du théorème 3.1 du chapitre 2 de [KL81]. On a ainsi une garantie de pouvoir fabriquer par produit des $w_{a,b}$ des unités modulaires s'annulant sur des orbites choisies.

On a plus précisément le résultat suivant, dérivé de la proposition 4.7 de [BP11b] qui découle elle-même après un certain travail technique des résultats précédents (le tout étant de bien choisir les produits des $w_{a,b}$).

Proposition III.3.22. *Soit Σ une partie propre de $\mathcal{C}(G,K)$ de cardinal s. Il existe une unité modulaire w définie sur K s'annulant sur toutes les orbites de Σ et $B \leq s^{s/2+1}(|G'|N^2)^{s-1})$ telle que :*

(a) Il existe un entier algébrique λ qui est le produit d'au plus $B|G'|$ facteurs de la forme $(1-\zeta_{N'})^{12N}$ avec $N'|N$, tel que w et λw^{-1} sont entières sur $\mathbb{Z}[j]$.

(b) Pour toute pointe c de X_G d'ordre de ramification e_c, toute place $v \in M_K$ finie et tout point P se réduisant en c modulo v :

$$|\log|w(P)|_v + \frac{\mathrm{ord}_c\, w}{e_c}\log|j(P)|_v| \leq B|G'|c_v.$$

(c) Pour toute place archimédienne v et tout point P, on a

$$|\log|w(P)|_v| \leq B|G'|N\log(|j(P)|_v + 2400) + B|G'|c_v.$$

Remarque III.3.7. L'apparition de $|G'|$ dans les bornes ci-dessus est due à la construction des produits $w_{a,b}$ à partir des $u_{a,b}$, et la constante B provient de la façon dont on doit faire les produits de ces $w_{a,b}$ pour s'assurer que les w s'annulent seulement là où on le souhaite. C'est la partie la moins explicite du résultat ci-dessus, qu'on peut espérer récupérer par d'autres outils pour certaines familles précises de courbes modulaires (pour $X_0(p)$ par exemple, on a la proposition I.7.2).

Nous allons maintenant pouvoir appliquer le coeur de la méthode de Runge et en déduire une borne sur la hauteur du j-invariant.

Soient G, K et G' comme notés ci-dessus. Soit P un point de $X_G(K)$ ayant bonne réduction sauf en les places $v \in S$ avec $|S| < |\mathcal{C}(G,K)|$. On note S_1 l'ensemble des places pour lesquelles $|j(P)|_v > 1$ (si v est non-archimédienne) ou $|j(P)[_v > 3500$ (si v est archimédienne), S_2 l'ensemble des places archimédiennes de $M_K \backslash S_1$ et S_3 le reste des places de M_K. D'après la proposition III.3.22, comme $|S_1| < |\mathcal{C}(G,K)|$, il existe une unité modulaire w s'annulant en toute pointe c dès que $P \in \Omega_{c,v}$ pour une certaine place $v \in S_1$.

On applique alors la formule du produit à $w(P)$, qui nous donne

$$0 = \sum_{v\in S_1} n_v \log|w(P)|_v + \sum_{v\in S_2} n_v \log|w(P)|_v + \sum_{v\in S_3} n_v \log|w(P)|_v.$$

Pour les places v de S_1, comme $P \in \Omega_{c_v,v}$ pour une certaine pointe c_v en laquelle s'annule w, on a d'après la proposition III.3.22 (b)

$$\log|w(P)|_v \leq B|G'|c_v - \frac{\mathrm{ord}_c\, w}{e_c}\log|j(P)|_v = B|G'|c_v - \frac{\mathrm{ord}_c\, w}{e_c}\log^+|j(P)|.$$

Pour les places v de S_2, on a d'après la proposition III.3.22 (c)

$$\log|w(P)|_v \leq B|G'|N\log(|j(P)|_v + 2400) + B|G'|c_v \leq B|G'|(N\log(5900) + c_v).$$

Enfin, pour les places v de S_3, on a par hypothèse $\log|w(P)|_v \leq 0$, d'où

$$0 \leq \sum_{v\in S_1} n_v(B|G'|c_v - \frac{\mathrm{ord}_c\, w}{e_c}\log^+|j(P)|_v) + \sum_{v\in S_2} n_v B|G'|(N\log(5900) + c_v).$$

En conséquence,

$$\sum_{v\in S_1} n_v \log^+ |j(P)|_v \leq N \sum_{v\in S_1} n_v B|G'|c_v + \sum_{v\in S_2} n_v B|G'|(N\log(5900)+c_v)$$

et par hypothèse, on a

$$\sum_{v\in S_2} n_v \log^+ |j(P)|_v \leq \sum_{v\in S_2} n_v \log(3500) \leq [K:\mathbb{Q}]\log(3500)$$

et

$$\sum_{v\in S_3} n_v \log^+ |j(P)|_v = 0,$$

d'où une borne absolue sur $h(j(P))$. Les calculs légèrement plus fins de [BP11b] aboutissent au théorème suivant.

Théorème III.6 (Runge pour les courbes modulaires, théorème 1.2 de [BP11b])**.**

Soit K un corps de nombres, $N \geq 1$ et G un sous-groupe de $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ tel que la courbe modulaire $X_G = X(N)/(G\cap \mathrm{SL}_2(\ /N\mathbb{Z}))$ est définie sur K une courbe modulaire. On note, pour toute extension L de K, $\mathcal{C}(G,L)$ l'ensemble des orbites des pointes de X_G. Alors, pour toute ensemble de places S de M_L contenant M_L^∞ tel que $s = |S| < |\mathcal{C}(G,L)|$ (c'est la condition de Runge), pour tout point $P \in X_G(L)$ tel que $j(P) \in \mathcal{O}_{L,S}$, on a

$$h_j(P) \leq 36 s^{s/2+1}(N^2|G|/2)^s \log(2N).$$

En particulier, la réunion de tous les ensembles de points ainsi obtenus est de hauteur h_j bornée par la fonction de s ci-dessus appliquée au nombre de pointes de X_G.

Remarque III.3.8. Dans le cas de $X_0(p)$, la méthode de Runge est appliquée dans la section I.7, et on obtient (car la situation est particulièrement favorable) une borne bien meilleure dans ce cas (proposition I.7.7). C'est aussi le cas pour $X_{\mathrm{split}}(p)$, celui-ci étant traité dans [BP11a].

III.4 Généralisation aux variétés de dimension quelconque

Commençons par la proposition suivante, clé de la preuve d'un théorème « à la Runge » en dimension supérieure.

Proposition III.4.1. *Soit K un corps de nombres. On utilise les notations de la définition III.1.3.*

Soit $X_K \subset \mathbb{P}^n_K$ une variété normale projective et $\phi_1, \cdots, \phi_r$ des fonctions rationnelles de $K(X)$. On note Y la sous-variété fermée de X constituée par l'intersection des supports des diviseurs des pôles des ϕ_i, et $g_1, \cdots, g_s$ des générateurs de l'idéal homogène de $K[X_0, \cdots, X_n]$ correspondant.

Pour tout $j \in \{1, \cdots, s\}$ et $i \in \{0, \cdots, n\}$, soit $g_{i,j} =: g_j \circ \varphi_i$ qui est une fonction régulière sur U_i. Alors, il existe une M_K-constante $(c_v)_{v\in M_K}$ telle que pour tout $v \in M_K$, tout $i \in \{0, \cdots, n\}$ et tout $(P,w) \in E_{i,v}$ avec $P \in X(\overline{K})$,

$$\min_{1\leq \ell \leq r} \log|\phi_\ell(P)|_w \leq c_v \quad \textit{ou} \quad \max_{1\leq j\leq s} \log|g_{i,j}(P)|_w < 0.$$

Remarque III.4.1. Dans le cas où $Y = 0$, on retrouve exactement le lemme 5 de [Lev08], c'est-à-dire que pour tout point P de $X(\overline{K})$ et toute place, une des valeurs de ϕ_ℓ est « assez petite ». Il est à noter que ce lemme 5 suppose que X est lisse, mais en fait la normalité de X suffit. Un point technique de cette preuve est qu'on peut se permettre de parler de générateurs de Y en tant que sous-variété de $\mathbb{P}^n$ (et non pas de X), mais on les considère ensuite comme des fonctions régulières sur les cartes affines de X obtenues grâce aux ouverts de coordonnées, et comme celle-ci est normale, on peut utiliser un Nullstellensatz « restreint à ces cartes affines ».

D'autre part, grâce à la remarque III.1.3 (*b*) (et en reprenant la preuve ci-dessous), on peut remplacer le cas

$$\max_{1\leq j\leq s}\log|g_{i,j}(P)|_w<0$$

par

$$\max_{1\leq j\leq s}\log|g_{i,j}(P)|_w<c_{0,v}$$

pour un choix de M_K-constante $(c_{0,v})_{v\in M_K}$, et alors la M_K-constante $(c_v)_{v\in M_K}$ obtenue par la proposition dépend de ce choix.

Démonstration. Pour tout $i\in\{0,\cdots,n\}$, on note $X_i=X\cap U_i$ qui est une variété affine normale sur K, et $\phi_{\ell,i}$ la restriction de ϕ_ℓ à X_i, qui est donc une fonction rationnelle sur X_i. Comme X est normale, on peut définir $(\phi_{\ell,i})_0$ le diviseur de Weil des zéros de $\phi_{\ell,i}$ sur X_i. A ce diviseur de Weil (positif), on associe l'idéal $I_{\ell,i}$ de $K[X_i]$ constitué des fonctions régulières h sur X_i telles que $\mathrm{div}(h)\geq(\phi_{\ell,i})_0$. Comme cet idéal est de type fini sur K, on en choisit des générateurs $h_{\ell,i,1},\cdots,h_{\ell,i,j_{\ell,i}}$. Les fonctions $h_{\ell,i,j}/\phi_{\ell,i}$ sont régulières sur X_i et $\mathrm{div}(h_{\ell,i,j}/\phi_{\ell,i})\geq(\phi_{\ell,i})_\infty$ où ce dernier est le diviseur des pôles de $\phi_{\ell,i}$, mais par construction de $I_{\ell,i}$, le minimum (diviseur de Weil premier par diviseur de Weil premier) des $\mathrm{div}(h_{\ell,i,j})$ est exactement $(\phi_{\ell,i})_0$: en effet, comme ce sont des générateurs de $I_{\ell,i}$, ce minimum est également le minimum pour tout $h\in I_{\ell,i}$, or pour chaque famille finie de diviseurs premiers de Weil distincts $D_1,\cdots,D_r,D'$ de X_i, il existe une uniformisante h pour D' d'ordre 0 pour chacun des $D_1,\cdots,D_r$ sinon l'idéal premier associé à D' serait inclus dans l'union finie des autres, ce qui est impossible car ils sont tous de hauteur 1. Ceci permet de fabriquer, pour tout diviseur premier D' de X_i n'appartenant pas au support de $(\phi_{\ell,i})_0$ une fonction h de $I_{\ell,i}$ d'ordre 0 le long de D' (et du bon ordre pour chaque D' dans le support de $(\phi_{\ell,i})_0$), donc ce minimum est bien $(\phi_{\ell,i})_0$. En conséquence, le minimum des diviseurs des $h_{\ell,i,j}/\phi_{\ell,i}$ est exactement $(\phi_{\ell,i})_\infty$, et donc les seuls zéros communs dans $\overline{K}$ de ces fonctions régulières sur X_i appartiennent au support de $(\phi_{\ell,i})_\infty$.

Ainsi, à i fixé et pour tous indices j,ℓ, les fonctions régulières $h_{\ell,i,j}/\phi_\ell$ sur X_i (variété affine fermée de U_i) ont comme seuls zéros communs dans $\overline{K}$ les éléments de $Y(\overline{K})$. D'après la proposition III.1.7 appliquée à chaque X_i puis regroupée (en prenant le maximum des M_K-constantes obtenues), il existe donc une M_K-constante $(c'_v)_{v\in M_K}$ telle que pour tout $v\in M_K$, tout $i\in\{0,\cdots,n\}$ et tout $(P,w)\in E_{i,v}$, on a

$$\max_{(\ell,j)}\log\left|\frac{h_{\ell,i,j}}{\phi_\ell}(P)\right|_w\geq-c'_v\quad\text{ou}\quad\max_{1\leq j\leq s}\log|g_{i,j}(P)|_w<0.$$

Or, il existe une deuxième M_K-constante $(c''_v)_{v\in M_K}$ telle que $\log|h_{\ell,i,j}(P)|_w\leq c''_v$ pour tout (ℓ,i,j) et tout $(P,w)\in E_{i,v}$ car $E_{i,v}$ est affinement M_K-bornée dans U_i. Si on n'est pas dans le deuxième cas de la dichotomie (remarquons qu'alors $\phi_\ell(P)$ est bien défini pour au moins un indice ℓ), il existe donc un triplet (ℓ,i,j) tel que

$$-c'_v\leq\log|h_{\ell,i,j}(P)|_w-\log|\phi_\ell(P)|_w\leq c''_v-\log|\phi_\ell(P)|_w$$

d'où

$$\log|\phi_\ell(P)|_w\leq c'_v+c''_v,$$

ainsi la M_K-constante $(c_v)_{v\in M_K}=(c'_v+c''_v)_{v\in M_K}$ satisfait l'énoncé de la proposition. □

Remarque III.4.2. Cette preuve généralise le résultat-clé de la preuve du théorème III.4. L'idée est la même, à ceci près qu'il faut prêter une plus grande attention à l'intersection des diviseurs de pôles (qui peut ne pas être vide) et à celle des diviseurs de zéros. Pour ce passage à la dimension supérieure, deux difficultés s'ajoutent à la situation des courbes : la première est que les diviseurs effectifs stricts ne sont pas forcément amples (ce qu'on utilisait implicitement avec Riemann-Roch), et la seconde est qu'on n'a pas encore fait la transition entre points entiers et norme v-adiques des évaluations pour certaines fonctions. En fait, contrairement au cas des courbes, nous allons d'abord définir les points entiers du point de vue de certains modèles entiers, et ensuite traduire ceci en termes de fonctions.

Définition III.4.2 (Points entiers sur un schéma projectif)**.** Soit K un corps de nombres et S_0 un ensemble de places contenant M_K^∞. Soit X un schéma projectif sur $\mathcal{O}_{K,S_0}$ et Y un sous-$\mathcal{O}_{K,S_0}$-schéma fermé de X. Alors, pour tout corps de nombres L/K :

(a) Pour toute place w de L associée à un idéal premier $\mathfrak{P}$ non au-dessus de S_0, les *points (L,w)-entiers de $X\backslash Y$* sont les éléments de l'ensemble

$$(X\backslash Y)(\mathcal{O}_{L,w}) = \{P \in X(L)\backslash Y(L) \,|\, P_{\mathfrak{P}} \notin Y_{\mathfrak{P}}\}.$$

Pour tout ensemble de places S_L contenant les places au-dessus de S_0, les *points (L,S_L)-entiers de $X\backslash Y$* sont les points (L,w)-entiers pour toute place w de $M_L\backslash S_L$, on note $(X\backslash Y)(\mathcal{O}_{L,S_L})$ l'ensemble de ces points.

(b) Pour un entier $s_L \geq 1$, les *points (L,s_L)-entiers de $X\backslash Y$* sont les points appartenant à un des ensembles $(X\backslash Y)(\mathcal{O}_{L,S_L})$ avec $|S_L| = s_L$ (autrement dit ayant chacun réduction hors de Y sauf en au plus s_L places, dont les places au-dessus de S_0).

Pour un théorème de Runge en dimension supérieure, nous allons encore une fois utiliser une propriété de Northcott, mais pour ceci il faut obtenir des plongements de notre variété projective obtenus à partir de fonctions rationnelles à pôles prescrits. C'est la raison pour laquelle on a besoin de propriété d'amplitude ou de grosseur de nos diviseurs.

Proposition III.4.3 (Détermination des points entiers)**.** *Avec les notations de la définition III.4.2 :*

(a) *Si Y est un diviseur effectif de X ample sur $\mathcal{O}_{K,S_0}$, il existe des fonctions $\phi_1, \cdots, \phi_n$ de $K(X)$ dont le diviseur des pôles est un multiple de Y_K, telles que l'application $X_K\backslash Y_K \to \mathbb{P}^n$ définie par $x \mapsto (1 : \phi_1(x) : \cdots : \phi_n(x))$ se prolonge en un plongement projectif $\psi_K : X_K \to \mathbb{P}^n_K$, et que pour tout couple (L,S_L) avec S_L contenant les places au-dessus de S_0,*

$$\forall P \in (X\backslash Y)(L), \Big(P \in (X\backslash Y)(\mathcal{O}_{L,S_L}) \Longleftrightarrow \forall i \in \{1,\cdots,n\}, \phi_i(P) \in \mathcal{O}_{L,S_L}\Big).$$

(b) *Si Y est un diviseur effectif de X tel que Y_K est ample, il existe des fonctions $\phi_1, \cdots, \phi_n$ de $K(X)$ dont le diviseur des pôles est un multiple de Y_K, telles que l'application $X_K\backslash Y_K \to \mathbb{P}^n_K$ définie par $x \mapsto (1 : \phi_1(x) : \cdots : \phi_n(x))$ se prolonge en un plongement projectif $\psi : X_K \to \mathbb{P}^n_K$, et une M_K-constante $(c_v)_{v\in M_K}$ telle que pour toute extension L de K et toute valeur absolue w de M_L au-dessus de $v \in M_K\backslash S_0$,*

$$\forall P \in (X\backslash Y)(L), \Big(P \in (X\backslash Y)(\mathcal{O}_{L,w}) \Longrightarrow \forall i \in \{1,\cdots,n\}, \log|\phi_i(P)|_w \leq c_v\Big).$$

(c) *Si Y est un diviseur effectif de X gros sur $\mathcal{O}_{K,S_0}$, il existe un fermé de Zariski strict Z de X et des fonctions $\phi_1, \cdots, \phi_n$ de $K(X)$ dont le diviseur des pôles est un multiple de Y_K telles que l'application $\psi_K : X_K\backslash(Y_K \cup Z_K) \to \mathbb{P}^n_K$ définie par $x \mapsto (1 : \phi_1(x) : \cdots : \phi_n(x))$ est une immersion, et que pour tout couple (L,S_L) avec S_L contenant les places au-dessus de S_0 et tout $P \in X\backslash(Y \cup Z)(L)$,*

$$P \in (X\backslash(Y \cup Z))(\mathcal{O}_{L,S_L}) \Longrightarrow \forall i \in \{1,\cdots,n\}, \phi_i(P) \in \mathcal{O}_{L,S_L}.$$

Ce fermé est un fermé quelconque tel qu'on a un plongement projectif $X\backslash Z \to \mathbb{P}^n$ sur $\mathcal{O}_{K,S_0}$ induit par les sections d'une certaine puissance de Y.

(d) *Si Y est un diviseur effectif de X tel que Y_K est gros, il existe un fermé de Zariski strict Z_K de X_K et des fonctions $\phi_1, \cdots, \phi_n$ de $K(X)$ dont le diviseur des pôles est un multiple de Y_K telles que l'application $\psi_K : X_K\backslash Y_K \to \mathbb{P}^n$ définie par $x \mapsto (1 : \phi_1(x) : \cdots : \phi_n(x))$ est une immersion lorsque restreinte à $X_K\backslash(Y_K \cup Z_K)$, et il existe une M_K-constante $(c_v)_{v\in M_K}$ telle que pour toute extension L de K et toute valeur absolue w de M_L au-dessus de $v \in M_K\backslash S_0$,*

$$\forall P \in (X\backslash Y)(L), \Big(P \in (X\backslash Y)(\mathcal{O}_{L,\mathfrak{P}}) \Longrightarrow \forall i \in \{1,\cdots,n\}, \log|\phi_i(P)|_w \leq c_v\Big).$$

Nous allons prouver cette proposition avant d'expliquer plus longuement le but de ses différentes parties.

Démonstration.

(a) Si Y est ample sur $\mathcal{O}_{K,S_0}$, il existe un plongement projectif $\psi : X \hookrightarrow \mathbb{P}^n$ tel qu'en tant que sous-schéma fermé de X, le support de ce diviseur Y est exactement l'image inverse par ψ du sous-schéma fermé de $\psi(X)$ défini par son intersection avec l'hyperplan $x_0 = 0$. On fixe un tel plongement, et on note $\phi_1, \cdots, \phi_n$ les fonctions rationnelles $(x_j/x_0) \circ \psi$ de $K(X)$. Leurs diviseurs de pôles sont des multiples de Y, donc l'application définie dans l'énoncé se prolonge bien en ψ. De plus, pour tout point $P \in X(L)\backslash Y(L)$ et tout idéal premier $\mathfrak{P}$ dont la place associée w n'est pas au-dessus de S_0, le point $P_{\mathfrak{P}}$ appartient à $Y_{\mathfrak{P}}$ si et seulement si la réduction modulo $\mathfrak{P}$ de $\psi(P)$ dans $\mathbb{P}^n(k(\mathfrak{P}))$ a sa première coordonnée nulle, ce qui équivaut à dire qu'il existe $i \in \{1, \cdots, n\}$ tel que $|x_i/x_0(\psi(P))|_w > 1$. Ainsi, $P_{\mathfrak{P}} \notin Y_{\mathfrak{P}}$ si et seulement si pour tout $i \in \{1, \cdots, n\}$, $\phi_i(P) \in \mathcal{O}_{L,w}$, ce qui appliqué place par place prouve le (a).

(b) Par amplitude de Y_K, il existe un plongement projectif $\psi_K : X_K \hookrightarrow \mathbb{P}^n_K$ tel qu'en tant que fermé de X_K, le support du diviseur Y_K est exactement l'image inverse par ψ_K du fermé de $\psi_K(X_K)$ défini par son intersection avec l'hyperplan $x_0 = 0$. On note $\phi_1, \cdots, \phi_n$ les fonctions rationnelles $(x_i/x_0) \circ \psi$ de $K(X)$ pour $i \in \{1, \cdots, n\}$. Leurs diviseurs de pôles sur la variété X_K sont des multiples de Y_K, et l'application définie dans l'énoncé du (c) se prolonge bien en ψ_K. Maintenant, on choisit un plongement quelconque de X dans un $\mathbb{P}^m_{\mathcal{O}_{K,S_0}}$, et des générateurs $g_1, \cdots, g_s$ de l'idéal homogène de $\mathcal{O}_{K,S_0}[X_0, \cdots, X_m]$ associé au sous-$\mathcal{O}_{K,S_0}$-schéma fermé Y vu dans $\mathbb{P}^m_{\mathcal{O}_{K,S_0}}$. Par construction, les $g_1, \cdots, g_s$ engendrent également l'idéal homogène de $K[X_0, \cdots, X_m]$ associé au fermé Y_K de $\mathbb{P}^m_K$. On applique alors la proposition III.4.1 à chaque $\phi_i, i \in \{1, \cdots, n\}$ et aux $g_1, \cdots, g_s$, et on prend le maximum des n M_K-constantes obtenues, noté $(c_v)_{v \in M_K}$. Ainsi, pour tout $j \in \{0, \cdots, m\}$, toute place w sur $\overline{K}$ au-dessus de v et tout point P de $(X\backslash Y)(\overline{K})$ tel que $P \in E_{j,w}$ (définition III.1.3), on a

$$\max_{1 \leq i \leq n} \log |\phi_i(P)|_w \leq c_v \quad \text{ou} \quad \max_{1 \leq k \leq s} \log |g_k \circ \varphi_j(P)|_w < 0.$$

Or, d'après la proposition III.1.8, le deuxième choix de cette dichotomie signifie que P se réduit dans Y pour w. En conséquence, si $P \in (X\backslash Y)(\mathcal{O}_{L,w})$, pour tout $i \in \{1, \cdots, n\}$, on a $\log |\phi_i(P)|_w \leq c_v$.

(c) Par définition de la grosseur, il existe un fermé de Zariski Z de X tel qu'on a un plongement projectif $\psi : X\backslash Z \to \mathbb{P}^n_{\mathcal{O}_{K,S_0}}$ identifiant $Y\backslash Z$ à l'intersection de l'image de ψ avec l'hyperplan d'équation $x_0 = 0$. On reprend le cas de l'amplitude ci-dessus, en prenant en compte le fait que ψ n'est un plongement que de $X\backslash Z$, d'où l'ajout des points Z-entiers dans l'implication. Ce n'est pas une équivalence car ϕ_i est définie sur un point de $(X\backslash Z)(L)$ mais on ne peut pas prédire l'intégralité de $\phi_i(P)$ si P se réduit dans Z modulo $\mathfrak{P}$.

(d) On réutilise la preuve du (b), en prenant en compte le fait que la grosseur de Y_K exclut tous les points P appartenant à un certain fermé Z_K de X_K (plus précisément, la propriété appliquée à chaque ϕ_i reste vraie et on peut les évaluer les points hors de Y_K, la seule chose qu'on perd par rapport à l'amplitude est la propriété de plongement de ψ_K hors de Y_K). □

Avant de donner une intuition plus précise sur ce résultat grâce à la remarque qui va suivre, nous allons mettre un peu de vocabulaire sur les propriétés dans chacun des cas.

Définition III.4.4 (Fonctions rationnelles et points entiers)**.** Avec les notations de la définition III.4.2 :

(a) Une fonction rationnelle $\phi \in K(X)$ *voit les points $X\backslash Y$-entiers* si le diviseur de ϕ est à support dans Y et pour toute place w d'une extension finie L de K non au-dessus de S_0, si $P \in (X\backslash Y)(\mathcal{O}_{L,w})$, alors $\phi(P) \in \mathcal{O}_{L,w}$.

(b) Une fonction rationnelle $\phi \in K(X)$ *voit les points $X\backslash Y$-entiers à M_K-constante près* si le diviseur de ϕ est à support dans Y et s'il existe une M_K-constante $(c_v)_{v \in M_K}$ telle que pour toute place w d'une extension finie L de K au-dessus de $v \notin S_0$, si $P \in (X\backslash Y)(\mathcal{O}_{L,w})$, alors $\log |\phi(P)|_w \leq c_v$.

(c) Un ensemble de fonctions rationnelles $\phi_1, \cdots, \phi_n \in K(X)$ *caractérise les points* $X\backslash Y$*-entiers* si chacun d'entre elles voit les points $X\backslash Y$-entiers et que réciproquement, pour tout w non au-dessus de S_0, si un point $P \in (X\backslash Y)(L)$ a son image dans $\mathcal{O}_{L,w}$ pour chaque ϕ_i, alors $P \in (X\backslash Y)(\mathcal{O}_{L,w})$.

Remarque III.4.3. Cette caractérisation semble nécessaire pour mettre en valeur la différence entre l'intégralité au sens de [Lev08] (vis-à-vis de diviseurs quelconques) et celle qu'on souhaite manipuler ici, à savoir en termes de réduction par rapport à un modèle entier.

Le vocabulaire introduit dans la définition ci-dessus permet de résumer les résultats obtenus dans la proposition III.4.3 : sous les hypothèses du (a), on obtient des fonctions rationnelles qui caractérisent les points $X\backslash Y$-entiers (et induisent un plongement projectif sur K). Sous les hypothèses du (b), on a encore un plongement projectif sur K mais les fonctions rationnelles voient seulement les points $X\backslash Y$-entiers à M_K-constante près. Sous les hypothèses du (c), le fermé introduit par la grosseur empêche de caractériser tous les points entiers, et on en voit donc seulement certains. Enfin, pour le (d), les fonctions rationnelles voient les points entiers de $X\backslash Y$ à M_K-constante près, sauf ceux de $Z(\overline{K})$.

Nous n'allons pas plus tard utiliser les caractérisations (a) et (c), qui supposent l'amplitude (resp. la grosseur) sur le schéma de base, car ces hypothèses sont difficiles à obtenir, mais nous avons préféré les donner pour expliquer la nuance avec leurs contreparties sur la fibre générique.

Pour mieux se convaincre qu'on n'a pas besoin de l'amplitude sur tout le schéma mais seulement en la fibre générique sur le (b), une approche analytique est peut-être préférable : l'idée est qu'une fonction rationnelle ϕ est w-grande seulement près du support de son diviseur de pôles. Dans le cas d'une variété projective, on peut identifier une certaine w-proximité à un diviseur de pôles à une réduction dans ce diviseur, et la caractériser par la petitesse de fonctions auxiliaires caractérisant ce diviseur. Le vrai passage théorique à des points entiers en termes de schéma se joue donc non pas grâce à ϕ mais grâce à ces fonctions auxiliaires, et c'est l'essence des énoncés des propositions III.1.8 et III.4.1. Il est important de noter que pour la suite, l'objectif final est un résultat de borne en termes de hauteur projective via un plongement (ou presque un plongement) : c'est à ce stade que l'amplitude ou la grosseur des diviseurs sera cruciale pour en déduire la finitude. C'est également pour pouvoir utiliser ceci qu'on a besoin d'avoir des fonctions rationnelles qui « voient à M_K-constante près » les points entiers : sans cela, on ne peut pas borner les hauteurs impliquées, car le point essentiel de ces fonctions rationnelles est de s'abstraire du découpage de $X(\overline{K})$ en ces $E_{i,v}$ (définition III.1.3), contrairement aux générateurs des idéaux de définition des supports des pôles.

Enfin, on peut réinterpréter des résultats précédents pour les courbes à la lueur de ce vocabulaire : l'essentiel de la traduction entre le théorème III.2 et le théorème III.4 consiste à trouver une fonction rationnelle ϕ qui voit à M_K-constante près les points entiers de $\mathcal{C}\backslash\mathcal{P}$, et pour les courbes modulaire (section III.3), la fonction j a l'avantage de caractériser à elle toute seule les points $\mathcal{C}\backslash\mathcal{P}$ lorsque $\mathcal{P}$ est le diviseur des pôles ([Sil09], Proposition VII.5.5).

Nous pouvons maintenant présenter un théorème de Runge en dimension supérieure, qui est une généralisation technique du théorème 4 ((b) et (c)) prouvé dans [Lev08], en prenant en compte la remarque III.4.3. L'argument que nous présentons ci-dessous permet l'uniformité la plus grande possible, en particulier en les extensions L vérifiant la condition de Runge (dans la preuve originale de [Lev08], la famille de plongements utilisée dépend de l'extension de K choisie).

Théorème III.7 (Runge en dimension supérieure)**.** *Soit K un corps de nombres et S_0 un ensemble fini de places de K contenant M_K^∞. Soit $\mathcal{O}$ la clôture intégrale de $\mathcal{O}_{K,S_0}$ dans une certaine extension finie K' de K.*

Soit X un schéma normal projectif sur $\mathcal{O}_{K,S_0}$ et $D_1, \cdots, D_r$ des diviseurs de Cartier effectifs sur $X_{\mathcal{O}}$ dont l'union des supports $D_{\mathcal{O}}$ est l'extension à $\mathcal{O}$ d'un certain sous-$\mathcal{O}_{K,S_0}$-schéma fermé D de X, et tels que $\mathrm{Gal}(K'/K)$ *permute les $(D_i)_{K'}$. Supposons que l'intersection de n'importe quelles $(m+1)$ fibres géométriques $(D_i)_{\overline{K}}$ des diviseurs est vide pour un certain entier $m \in \mathbb{N}_{>0}$. Pour tout corps de nombres L/K, on note r_L le nombre d'orbites galoisiennes de l'ensemble des $(D_i)_{K'}$ par* $\mathrm{Gal}(K'L/L)$*. Alors :*

(a) Si $(D_1)_{\overline{K}}, \cdots, (D_r)_{\overline{K}}$ sont amples, la réunion de tous les ensembles de points (L, s_L)-entiers de $(X \backslash D)$ tels que

$$ms_L < r_L$$

(condition de Runge multidimensionnelle) est un ensemble fini.

(b) Si $(D_1)_{\overline{K}}, \cdots, (D_r)_{\overline{K}}$ sont gros, il existe un fermé de Zariski strict Z_K de X_K telle que la réunion de tous les ensembles de points (L, s_L)-entiers de $(X \backslash D)$ n'appartenant pas à $Z(L)$ est un ensemble fini. Le fermé $Z_{\overline{K}}$ est l'union des fermés $Z_{i,\overline{K}}$ tels que les plongements (sur $\overline{K}$) obtenus par la proposition III.4.3 (d) sont définis respectivement en-dehors de $Z_{1,\overline{K}}, \cdots, Z_{r,\overline{K}}$.

Remarque III.4.4. Le passage par K' et $\mathcal{O}$ est une nécessité technique pour prendre en compte le fait que les diviseurs en jeu n'ont pas besoin d'être des sous-schémas fermés de X sur K, mais sur une extension de ses scalaires, du moment que leur union de supports provient d'un sous-schéma fermé de X. Nous n'avons pas ici, contrairement au résultat de Levin, supposé les $(D_i)_{K'}$ irréductibles, mais pour pouvoir employer la même preuve, il faut supposer que $\mathrm{Gal}(K'/K)$ permute les $(D_i)_{K'}$, ce qui est automatique si les $(D_i)_{K'}$ sont irréductibles, comme $D_{K'}$ est stable par $\mathrm{Gal}(K'/K)$. La situation est la même que dans les hypothèses du théorème III.4, où il suffit que l'ensemble de points $\mathcal{P}$ soit stable par $\mathrm{Gal}(\overline{K}/K)$ et pas que chacun de ces points soit à valeurs dans K. Cette formulation permet donc d'éviter de faire l'extension des scalaires pour retrouver les diviseurs $D_1, \cdots, D_r$, ce qui nous empêcherait de considérer la condition de Runge seulement pour une extension de K' et non pas de K.

Pour le (b), la puissance du résultat dépend fortement de quels fermés de Zariski sont obtenus par grosseur de $(D_1)_{\overline{K}}, \cdots, (D_r)_{\overline{K}}$, c'est-à-dire de la géométrie particulière des $(D_i)_{\overline{K}}$.

Démonstration.

(a) Par hypothèse d'amplitude, pour tout $i \in \{1, \cdots, r\}$, d'après la proposition III.4.3 (b), il existe un plongement $\psi_i : X_{K'} \hookrightarrow \mathbb{P}^{n_i}_{K'}$ tel que le support de $(D_i)_{K'}$ est envoyé sur l'intersection de $\psi_i(X_{K'})$ avec l'hyperplan $x_0 = 0$, et que les fonctions coordonnées $\phi_{i,1}, \cdots, \phi_{i,n_i}$ « voient à $M_{K'}$-constante près » les points entiers de $X \backslash D_i$. On note $(c_w)_{w \in M_{K'}}$ le maximum des r $M_{K'}$-constantes produites par cette proposition, dont on déduit une M_K-constante $(c_v)_{v \in M_K}$ selon la remarque III.1.1 (a). De plus, si deux diviseurs $(D_i)_{K'}$ et $(D_{i'})_{K'}$ sont conjugués par un certain $\sigma \in \mathrm{Gal}(K'/K)$ (c'est-à-dire que $\sigma((D_i)_{K'}) = (D_{i'})_{K'}$, on peut choisir (et on le fait désormais) les plongements de sorte que ${}^{\sigma}\psi_i = \psi_{i'}$, et alors $n_i = n_{i'}$ et pour tout $j \in \{1, \cdots, n_i\}$, on a ${}^{\sigma}\phi_{i,j} = \phi_{i',j}$.

Soit maintenant $\mathcal{J}$ une partie de $\{(i,j),\ 1 \leq i \leq r,\ 1 \leq j \leq n_i\}$ telle que les fonctions $\phi_{i,j}$ avec $(i,j) \in \mathcal{J}$ sont sans pôle commun. À chaque telle partie $\mathcal{J}$, on peut associer une $M_{K'}$-constante grâce à la proposition III.4.1 (donc une M_K-constante par la remarque III.1.1 (a)), et on prend leur maximum à toutes, notée $(c'_v)_{v \in M_K}$. Soit maintenant L une extension de K, S_L un ensemble de places de M_L contenant les places au-dessus de S_0, et P un point (L, S_L)-entier de $X \backslash D$. On note $L' = K'L$ dans une clôture algébrique fixée de K. Alors, d'après la proposition III.4.3 (b), pour tout couple (i,j) et toute place w de $M_{L'}$ au-dessus de $v \notin S_0$, $\log |\phi_{i,j}(P)|_w \leq c_v$. De plus, par définition de m et de la M_K-constante $(c_v)_{v \in M_K}$, pour toute place $w \in M_{K'}$, l'ensemble des couples (i,j) tels que $\log |\phi_{i,j}(P)|_w > c'_v$ ne contient pas plus de m indices i différents. Ensuite, si $\sigma(D_i) = D'_i$ pour un certain $\sigma \in \mathrm{Gal}(L'/L)$, vu notre construction des plongements, pour toute place w de $M_{L'}$, et tout $j \in \{1, \cdots, n_i\}$, on a, comme P est défini sur L :

$$|\phi_{i',j}(P)|_w = |\phi_{i',j}(\sigma(P))|_w = |({}^{\sigma}\phi_{i,j})(\sigma(P))|_w = |\sigma(\phi_{i,j}(P))|_w = |\phi_{i,j}(P)|_{\sigma^{-1}(w)}.$$

Ceci prouve que pour toute place v de M_L, l'ensemble des $(D_i)_{K'}$ tels qu'il existe $j \in \{1, \cdots, n_i\}$ et une place w de $M_{L'}$ au-dessus de v telle que $\log |\phi_{i,j}(P)|_w > \max(c_v, c'_v)$ est stable par $\mathrm{Gal}(L'/L)$.

Par définition de m, pour chaque place v de M_L, il existe donc au plus m indices i tels que pour un des $j \in \{1, \cdots, n_i\}$, $\log |\phi_{i',j}(P)|_w > \max(c_v, c'_v)$ pour au moins une des places w de L' au-dessus de v (l'intérêt de ce raisonnement est de ne pas perdre plus d'indices à cause de l'augmentation du nombre de places entre L et L'). En supposant que la condition de Runge est vérifiée pour s_L et le nombre d'orbites r_L des $(D_i)_{L'}$ pour $\mathrm{Gal}(L'/L)$, il reste donc un indice i tel

que pour tout $j \in \{1, \cdots, n_i\}$ et toute place w de $M_{L'}$, on a

$$\log |\phi_{i,j}(P)|_w \leq \max(c_v, c'_v).$$

En résumé, en notant $(c''_v)_{v \in M_K}$ le maximum des deux M_K-constantes $(c_v)_{v \in M_K}$ (fournie par le fait que les fonctions rationnelles utilisées « voient à M_K-constante près » les points entiers) et $(c'_v)_{v \in M_K}$ (fournie par le fait que les fonctions rationnelles utilisées, ayant des pôles d'intersection vide, ne peuvent pas être toutes grandes en même temps), on obtient

$$h_{\psi_i}(P) \leq \frac{1}{[L' : \mathbb{Q}]} \sum_{w \in M_{L'}} n_w c''_w \leq \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} n_v c''_v,$$

d'après la remarque III.1.1 (a). L'ensemble des tous les points (L, S_L)-entiers de $X \backslash D$ pour une certaine paire (L, S_L) vérifiant la condition de Runge est donc inclus dans l'union finie de r ensembles dont chacun est constitué de points de degré borné (car on doit avoir $[L : \mathbb{Q}] < 2r$ pour pouvoir vérifier la condition de Runge) et de hauteur absolument bornée via un plongement projectif, donc fini par la propriété de Northcott.

(b) Dans ce cas, on reprend quasiment mot à mot la preuve précédente, en n'oubliant pas de considérer pour chaque ψ_i qu'il n'est un plongement qu'en-dehors d'un fermé $Z_{i,K'}$ et en utilisant la proposition III.4.3 (d). On obtient par exactement le même procédé la finitude de l'ensemble considéré. □

Dans de nombreux cas rencontrés en dimension supérieure (y compris ceux de la section III.5), on peut avoir affaire à des diviseurs dont l'intersection n'est jamais vide. C'est en fait une tendance naturelle de la dimension supérieure : pour appliquer Runge comme ci-dessus, on a besoin de diviseurs amples (comme « bords ») par rapport auxquels définir l'intégralité. Or d'après les critères numériques d'amplitude, les diviseurs amples s'intersectent automatiquement entre eux, ce qui fait donc monter le nombre m d'intersections nécessaire pour obtenir l'ensemble vide et peut rendre la condition de Runge trop exigeante.

On voudra donc considérer le cas où l'intersection de m diviseurs n'est certes pas vide mais suffisamment petite, et supposer que notre point ne se réduira jamais en cette intersection pour pouvoir appliquer la méthode de Runge.

On aura cependant besoin d'une telle hypothèse en toutes les places, y compris archimédiennes (donc sans interprétation naturelle en termes de réduction), ce qui motive la définition suivante.

Définition III.4.5 (Points parfaitement $X \backslash Y$-entiers)**.** Soit X un schéma projectif sur $\mathcal{O}_{K,S_0}$ et Y un sous-$\mathcal{O}_{K,S_0}$-schéma fermé de X. On reprend les notations de la définition III.1.3.

Soit un plongement projectif $X \subset \mathbb{P}^n_{\mathcal{O}_{K,S_0}}$ et un choix de générateurs $g_1, \cdots, g_s$ de l'idéal homogène de $\mathcal{O}_{K,S}[X_0, \cdots, X_n]$ définissant Y. Alors, avec ces choix :

Pour toute extension L de K, un point $P \in X(L)$ est *parfaitement $X \backslash Y$-entier* pour $g_1, \cdots, g_s$ si pour *toute* place v de M_L et tout $i \in \{0, \cdots, n\}$ tel que $P \in E_{i,v}$, il existe $j \in \{1, \cdots, r\}$ tel que

$$|g_j \circ \varphi_i(P)|_v \geq 1.$$

Plus généralement, pour une M_K-constante $(c_{0,v})_{v \in M_K}$, un point $P \in X(L)$ est *quasi-$X \backslash Y$-entier pour $g_1, \cdots, g_s$ et $(c_{0,v})_{v \in M_K}$* si pour *toute* place v de M_L et tout $i \in \{0, \cdots, n\}$ tel que $P \in E_{i,v}$, il existe $j \in \{1, \cdots, r\}$ tel que

$$\log |g_j \circ \varphi_i(P)|_v \geq c_{0,v}.$$

Remarque III.4.5. Tout d'abord, en toute place v de M_L non au-dessus de S_0, la parfaite intégralité équivaut à la non-réduction dans Y d'après la proposition III.1.8. L'utilité de ce renforcement de la notion d'intégralité est donc d'assurer qu'en place v de S_0, P soit v-adiquement loin de Y : dans le cas le plus simple où $K = \mathbb{Q}$ et $v = \infty$, on veut donc en plus de l'intégralité habituelle que P soit hors d'un voisinage de Y, qui aura donc une forme de boule si Y est un

point, de tube si Y est une courbe, de nappe épaissie si Y est une surface, et ainsi de suite. Nous allons dans tous les cas, par la suite, parler de *voisinage tubulaire de Y* pour cette notion.

On pourra avoir besoin d'affaiblir cette condition d'intégralité (notamment pour les évaluations en des places archimédiennes, où le choix de la M_K-constante nulle n'a pas de signification particulière), d'où la définition de quasi-M_K-intégralité. Pour la suite, les résultats valables pour des points parfaitement entiers le seront également pour les points quasi-M_K-entiers (à choix de M_K-constante fixé).

Cette définition implique trivialement la finitude dans le cas d'un diviseur ample, comme nous allons le montrer tout de suite.

Proposition III.4.6. *Avec les notations précédentes, supposons de plus que Y est un diviseur effectif ample sur $\mathcal{O}_{K,S_0}$ et choisissons une M_K-constante $(c_{0,v})_{v\in M_K}$. Alors, pour un (autre) plongement $\psi : X \to \mathbb{P}^m$ associé à Y via la proposition III.4.3 (a), et les fonctions rationnelles associées $\phi_1, \cdots, \phi_m \in K(X)$, il existe une M_K-constante $(c_v)_{v\in M_K}$ telle que pour tout point $P \in X(L)$ quasi $X\backslash Y$-entier pour $g_1, \cdots, g_s$ et $(c_{0,v})_{v\in M_K}$, tout $i \in \{1, \cdots, m\}$ et toute extension w de v à M_L,*

$$|\log \phi_i(P)|_w \leq c_v.$$

En conséquence, la hauteur h_ψ de l'ensemble des points quasi $X\backslash Y$-entiers pour $g_1, \cdots, g_s$ et $(c_{0,v})_{v\in M_K}$ est bornée, donc si on borne aussi leur degré, ils forment un ensemble fini.

Démonstration. Il suffit d'appliquer pour chaque ϕ_i la proposition III.4.1 et la remarque III.4.1, et on n'est jamais dans le deuxième cas de la dichotomie car on a supposé la quasi-$X\backslash Y$-intégralité. Remarquons que dans cette proposition, on n'a pas besoin pour les ϕ_i de connaître plus que leur diviseur de pôles. En prenant $(c_v)_{v\in M_K}$ le maximum des M_K-constantes obtenues, on a la proposition avec cette M_K-constante. □

Le but de Runge étant précisément de réussir à borner également les contributions à la hauteur en les places au-dessus de S_0, on va évidemment chercher à ce que l'hypothèse de quasi-$X\backslash Y$-intégralité soit limitée au strict minimum. Sous les hypothèses précédentes, on peut maintenant formuler une généralisation des théorèmes III.4 et III.7, fonctionnant plus souplement en fonction des intersections de diviseur, mais nécessitant cette notion.

Théorème III.8 (Théorème de Runge-Bombieri tubulaire)**.** *Soit K un corps de nombres, S_0 un ensemble fini de places de K contenant M_K^∞. Soit $\mathcal{O}$ la clôture intégrale de $\mathcal{O}_{K,S_0}$ dans une extension finie K' de K.*

Soit X un schéma normal projectif sur $\mathcal{O}_{K,S_0}$ et $D_1, \cdots, D_r$ des diviseurs de Cartier effectifs de $X_\mathcal{O}$ dont l'union des supports $D_\mathcal{O}$ est l'extension à $\mathcal{O}$ d'un sous-$\mathcal{O}_{K,S_0}$-schéma fermé D de X et tels que $\mathrm{Gal}(K'/K)$ permute les $(D_i)_{K'}$. Supposons qu'il existe un sous-$\mathcal{O}_{K,S_0}$-schéma fermé Y de X tel que l'intersection de n'importe quelles $(m+1)$ fibres génériques $(D_i)_{\overline{K}}$ est incluse dans $Y_{\overline{K}}$, et on choisit des générateurs homogènes $g_1, \cdots, g_t$ de l'idéal de définition de Y pour un certain plongement projectif $X \subset \mathbb{P}^n_{\mathcal{O}_{K,S_0}}$. Pour tout corps de nombres L/K, on note r_L le nombre d'orbites de l'ensemble des $(D_i)_{K'}$ par $\mathrm{Gal}(K'L/L)$. Alors :

(a) Si les fibres géométriques $(D_1)_{\overline{K}}, \cdots, (D_r)_{\overline{K}}$ sont amples, la réunion de tous les ensembles de points (L, s_L)-entiers de $X\backslash D$ et $X\backslash Y$-parfaitement entiers pour $g_1, \cdots, g_t$ tels que

$$ms_L < r_L$$

(condition de Runge multidimensionnelle) est un ensemble fini.

(b) Si les fibres géométriques $(D_1)_{\overline{K}}, \cdots, (D_r)_{\overline{K}}$ sont grosses, il existe un fermé de Zariski strict Z_K de X_K tel que la réunion de tous les ensembles de points (L, s_L)-entiers de $X\backslash D$, n'appartenant pas à $Z_K(\overline{K})$ et $X\backslash Y$-parfaitement entiers pour $g_1, \cdots, g_t$ (où (L, s_L) vérifie la condition de Runge multidimensionnelle) est un ensemble fini. Le fermé $Z_{\overline{K}}$ est l'union des fermés $Z_{i,\overline{K}}$ tels que les plongements sur $\overline{K}$ obtenus par grosseur des fibres génériques $(D_1)_{\overline{K}}, \cdots, (D_r)_{\overline{K}}$ sont définis respectivement en-dehors de $Z_{1,\overline{K}}, \cdots, Z_{r,\overline{K}}$.

Démonstration. La preuve du (a) et du (b) sont presque identiques à leurs pendants du théorème III.7 : grâce à la dichotomie de la proposition III.4.1 (dont le deuxième cas est automatiquement exclu par intégralité parfaite), il existe une $M_{K'}$-constante telle qu'on a, pour chaque place $w \in M_{L'}$, au plus m indices $i \in \{1, \cdots, r\}$ tels qu'il existe j vérifiant $\log |\phi_{i,j}(P)|_w > c_v$. De là, la preuve se poursuit telle que pour le théorème 13 aussi bien pour le (a) que pour le (b) sans changer un mot, et on aboutit donc au fait que l'ensemble de tous ces points est inclus dans une union finie d'ensembles de hauteur bornée dans un espace projectif, d'où la finitude.

□

Remarque III.4.6. Encore une fois, grâce aux remarques III.4.1 et III.4.5, le théorème III.8 est encore valide en remplaçant l'intégralité parfaite par la quasi-intégralité pour une M_K-constante fixée. Ceci permettra d'assouplir cette intégralité à notre convenance. Ainsi, on peut voir (en manipulant seulement la M_K-constante sur les places archimédiennes) ce théorème de Runge comme un résultat de concentration des points entiers (pour des paires vérifiant la condition de Runge) vers le domaine Y : autrement dit, à un nombre fini d'exceptions près, ces points entiers doivent être proche du fermé Y dans une certaine place archimédienne. Le théorème CLZ de [CLZ09], est de saveur similaire, avec les avantages et inconvénients suivants. Tout d'abord, il n'y a pas de condition sur la taille de l'ensemble de places S_L, pas de condition aussi stricte d'intersection que dans notre théorème, et pas non plus d'hypothèse supplémentaire d'intégralité parfaite (ou quasi-intégralité). Par contre, la finitude qu'il formule (bien que le domaine d'exclusion Y soit absolu) dépend de l'ensemble de places S_L (en particulier, il ne dit pas si la réunion de tous ces ensembles finis est finie ou non), et on ne sait rien dans le cas des diviseurs gros. Enfin, il semblerait que la preuve employée (passant par le théorème du sous-espace de Schmidt) est de nature non effective, contrairement à celle de notre théorème si on dispose des fonctions en jeu.

Par ailleurs, la formulation du choix de Y peut sembler curieuse (notamment parce qu'on pourrait envisager de l'appliquer à Y contenu dans une fibre fermée de X), mais il se trouve qu'elle reste valide même dans un tel cas de figure. En pratique, on pourra cependant penser à Y comme à l'adhérence de Zariski dans X d'un certain sous-schéma-fermé de X_K mais dans certains cas, on préférera définir directement Y pour éviter d'avoir à prouver que l'adhérence de Zariski de Y_K est bien Y.

Remarque III.4.7. Notre théorème de Runge tubulaire se présente dans sa formulation même comme une stratification d'énoncés « à la Runge » suivant la dimension de l'intersection des diviseurs en jeu. À un extrême de cette stratification, l'intersection en question est vide, et on retrouve alors le théorème III.7, et à l'autre, l'intersection est un diviseur (ample ou gros suivant le cas (a) ou (b)) d'où la finitude d'après la proposition III.4.6.

Il est à noter que cette stratification ne généralise les théorèmes préexistants qu'en dimension au moins égale à 2, car dans le cas des courbes, il n'y a pas de situation intermédiaire entre ces deux extrêmes, et les diviseurs strictement effectifs sont amples grâce au théorème de Riemann-Roch.

Cette généralisation où l'on autorise un lieu de concourance non vide des mauvais diviseurs est naturelle dès qu'on passe des courbes aux surfaces. Par exemple, si la fibre générique de X est de dimension 2, deux diviseurs amples sur X s'intersectent nécessairement, mais si dans notre problème on peut accepter l'intégralité parfaite par rapport à leurs points d'intersection, cela permet d'avoir $m = 1$ dans la condition de Runge si les diviseurs amples sont concourants en ces points d'intersection, au lieu de $m = 2$ ou plus.

III.5 Runge et les variétés modulaires de Siegel

Dans cette section, nous allons présenter diverses perspectives d'applications de la méthode de Runge en dimension supérieure aux variétés modulaires de Siegel, pour obtenir des résultats de finitude ayant une certaine signification modulaire, tout comme dans le cas des courbes modulaires.

Le paragraphe III.5.1 est consacré aux nombreux rappels et définitions sur les variétés modulaires de Siegel. Il contient de nombreux résultats non triviaux que nous avons extraits du chapitre VII de [Deb99] (pour les premières définitions, la polarisation et la structure de niveau dans le cas

complexe), puis de [Nam80] (surtout les chapitres 4,5 et 9), et enfin de [FC90] (chapitres IV et V) pour l'algébrisation des compactifications. Ensuite, nous donnerons les propriétés géométriques utiles pour envisager d'appliquer la méthode de Runge à ces variétés modulaires : celles-ci sont dispersées dans la littérature et pour certaines d'entre elles assez récentes, mais les références principales sont [Mum83], [Hul00] et [HW00]. Pour finir, nous donnerons une application concrète de la méthode de Runge pour les zéros de fonctions thêta, suivie de résultats conjecturaux pour certains diviseurs intéressants sur les variétés modulaires, en insistant particulièrement sur les nouvelles difficultés rencontrées par rapport au cas des courbes modulaires.

III.5.1 Rappels sur les $\mathcal{A}_g(n)$ et leurs compactifications

Commençons par construire les variétés modulaires de Siegel complexes en tant qu'espaces symétriques.

Définition III.5.1 (Demi-espace supérieur de Siegel, et groupe symplectique)**.**

Soit $g \geq 1$ un entier. On voit ici les éléments de $\mathbb{Z}^g, \mathbb{R}^g$ ou $\mathbb{C}^g$ comme des vecteurs lignes de taille g.

(a) Le *demi-espace supérieur de Siegel* d'ordre g, noté $\mathcal{H}_g$, est l'ensemble

$$\mathcal{H}_g := \{\tau \in M_g(\mathbb{C}) \,|\, {}^t\tau = \tau \text{ et } \operatorname{Im}\tau > 0\}$$

où $\operatorname{Im}\tau > 0$ signifie que $\operatorname{Im}\tau$ est définie positive en tant que matrice symétrique réelle. C'est un ouvert du $\mathbb{C}$-espace vectoriel des matrices symétriques complexes de taille g (muni de sa topologie naturelle), et en tant que tel, il est muni d'une structure de variété complexe. Pour tout $\tau \in \mathcal{H}_g$, on note A_τ le tore complexe $\mathbb{C}^g/\Lambda_\tau$ avec $\Lambda_\tau = \mathbb{Z}^g + \mathbb{Z}^g\tau$.

(b) Soit $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_{2g}(\mathbb{Z})$. Pour un anneau commutatif A, le *groupe symplectique de degré g sur A*, noté $\mathrm{Sp}_{2g}(A)$, est le sous-groupe de $\mathrm{GL}_{2g}(A)$ défini par

$$\mathrm{Sp}_{2g}(A) := \{M \in \mathrm{GL}_{2g}(A) \,|\, {}^tMJM = J\}.$$

Une matrice de $M_{2g}(A)$ est *symplectique* si elle appartient à ce groupe. Pour tout $n \geq 1$, le *groupe symplectique principal de degré g et de niveau n*, noté $\Gamma_g(n)$, est le sous-groupe de $\mathrm{Sp}_{2g}(\mathbb{Z})$ constitué des matrices congrues à l'identité modulo n. Pour tout $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{R})$ et tout $\tau \in \mathcal{H}_g$, on note

$$j_\gamma(\tau) := C\tau + D \in M_g(\mathbb{C}). \tag{III.5}$$

On définit également l'action à gauche de γ sur τ par

$$\gamma \cdot \tau := (A\tau + B)(C\tau + D)^{-1}. \tag{III.6}$$

Chaque élément de $\mathrm{Sp}_{2g}(\mathbb{R})$ agit alors par biholomorphismes sur $\mathcal{H}_g$, et j_γ est un cocycle pour cette action (voir la proposition VII.1.1 de [Deb99] pour les détails).

Définition III.5.2 (Tore complexe et polarisation)**.**

Soit $X = V/\Lambda$ un tore complexe, avec V un $\mathbb{C}$-espace vectoriel complexe et Λ un réseau de V.

(a) Une *polarisation de X* est une forme $\mathbb{R}$-bilinéaire alternée ω sur $V \times V$, à valeurs entières sur $\Lambda \times \Lambda$ et telle que pour tous $x, y \in V$, $\omega(ix, iy) = \omega(x, y)$ et $\omega(x, ix) > 0$ si x est non nul. Un tore complexe admettant une polarisation est une *variété abélienne complexe*.

(b) Avec les notations précédentes, il existe une $\mathbb{Z}$-base (dite adaptée) de Λ et des diviseurs élémentaires $d_1 | \cdots | d_g$ uniquement déterminés par ω tels que dans cette base, la matrice de ω est

$$\begin{pmatrix} 0 & D \\ -D & 0 \end{pmatrix}, \text{ où } D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_g \end{pmatrix}.$$

Le *degré de la polarisation* ω est alors le produit des entiers $d_1 \cdots d_g$, en particulier la polarisation est *principale* si ce degré est 1, c'est-à-dire si la matrice ci-dessus est J.

(c) Pour tout $n \geq 1$, et toute variété abélienne complexe principalement polarisée (X, ω) (ou même de polarisation de degré premier à n), le groupe $X[n]$ des points de n-torsion admet une forme bilinéaire w_n non dégénérée à valeurs dans μ_n le groupe des racines n-ièmes de l'unité dans $\mathbb{C}$, définie par $w_n(x,y) = e^{2i\pi n\omega(\tilde{x},\tilde{y})}$ où $\tilde{x}$ et $\tilde{y}$ sont des relèvements quelconques de x et y dans V. Cette forme bilinéaire alternée de $X[n]$ vers μ_n est appelée *l'accouplement de Weil sur* $X[n]$.

(d) Une *variété abélienne complexe principalement polarisée avec structure de niveau* n est la donnée d'une variété abélienne complexe $X = V/\Gamma$, d'une polarisation principale ω sur X, et d'un isomorphisme $\alpha_n : X[n] \to (\mathbb{Z}/n\mathbb{Z})^{2g}$ tel que pour la forme bilinéaire

$$W_n : \left| \begin{array}{ccc} (\mathbb{Z}/n\mathbb{Z})^{2g} \times (\mathbb{Z}/n\mathbb{Z})^{2g} & \longrightarrow & \mu_n \\ (X,Y) & \longmapsto & e^{\frac{2i\pi}{n} X J {}^t Y} \end{array} \right. ,$$

on a $\alpha_n^* W_n = w_n$ l'accouplement de Weil sur $X[n]$.

Toute $(\mathbb{Z}/n\mathbb{Z})$-base de $X[n]$ obtenue via cet isomorphisme α_n est dite *base symplectique de* $X[n]$ *(pour* ω*)*. Le choix d'un isomorphisme α_n équivaut au choix d'une base symplectique, et deux bases symplectiques sont liées par une matrice de passage appartenant à $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$.

(e) Deux variétés abéliennes complexes principalement polarisées (X, ω, α_n) et (X', ω', α'_n) avec structure de niveau n sont *isomorphes* s'il existe un isomorphisme d'espaces vectoriels $\phi : V \to V'$ (si $X = V/\Lambda$ et $X' = V/\Lambda'$) tel que $\phi(\Lambda) = \Lambda'$, $\phi^*\omega' = \omega$ et $\overline{\phi}^* \alpha'_n = \alpha_n$, où $\overline{\phi}$ est l'isomorphisme $X \cong X'$ induit par ϕ.

Remarque III.5.1. La donnée d'une polarisation principale sur X équivaut à la donnée de la première classe de Chern d'un fibré en droites ample L sur X tel que $h^0(X, L) = 1$, et on peut en déduire intrinsèquement une construction de l'accouplement de Weil et les notions d'isomorphismes de structures principales de niveau n (voir [Deb99], section VI.4). Nous avons opté pour les définitions ci-dessus pour leur caractère plus explicite, mais sur un corps quelconque, il faut effectivement définir les polarisations par classe de Chern, ou encore par isogénie entre A et sa variété abélienne duale $\widehat{A}$.

La proposition suivante explique le lien entre les deux définitions précédentes. Les résultats énoncés sont classiques, nous allons donc simplement en indiquer des références de preuves.

Définition-Proposition III.5.3.

Soit $g \in \mathbb{N}_{>0}$. On reprend les notations de la définition III.5.1.

(a) Pour tout $\tau \in \mathcal{H}_g$, la forme bilinéaire ω_τ sur $\mathbb{C}^g$ définie par

$$\omega_\tau(x,y) = \mathrm{Re}(x)\,\mathrm{Im}(\tau)^{-1\,t}\,\mathrm{Im}(y) - \mathrm{Re}(y)\tau^{-1}\,\mathrm{Im}(\tau)^{-1\,t}\,\mathrm{Im}(x)$$

est une polarisation principale du tore complexe A_τ, telle que pour tous $(i,j) \in \{1, \cdots, g\}^2$, pour $(e_1, \cdots, e_g)$ la base canonique de $\mathbb{C}^g$:

$$\omega_\tau(e_i, e_j) = 0 \quad \text{et} \quad \omega_\tau(e_i, e_j\tau) = \delta_{ij}.$$

(b) Pour tout $n \in \mathbb{N}_{>0}$, en notant $\pi : \mathbb{C}^g \to A_\tau$ la projection canonique, la famille $(\pi(e_1/n), \cdots, \pi(e_g/n), \pi(e_1\tau/n), \cdots, \pi(e_g\tau/n))$ est une base symplectique de $A_\tau[n]$ pour ω_τ et on note $\alpha_{\tau,n}$ la structure principale de niveau n associée.

(c) L'action de $\Gamma_g(1)$ sur $\mathcal{H}_g$ est proprement discontinue, c'est-à-dire que pour tout $\tau \in \mathcal{H}_g$, il existe un ouvert U de $\mathcal{H}_g$ contenant τ tel que $\{\gamma \in \Gamma_g(1) \,|\, \gamma \cdot U \cap U \neq \emptyset\}$ est fini. Pour tout $n \geq 3$, l'action de $\Gamma_g(n)$ est même librement discontinue.

(d) Pour tout $n \in \mathbb{N}_{>0}$, le quotient $A_g(n)_\mathbb{C} := \Gamma_g(n)\backslash\mathcal{H}_g$ est l'espace de modules des variétés abéliennes principalement polarisées munies d'une structure principale de niveau n, c'est-à-dire que l'application $\tau \mapsto (A_\tau, \omega_\tau, \alpha_{n,\tau})$ induit une bijection entre $\Gamma_g(n)\backslash\mathcal{H}_g$ et l'ensemble des variétés abéliennes principalement polarisées avec structure de niveau n à isomorphisme près. De plus, le quotient $A_g(n)_\mathbb{C}$ a une structure naturelle d'espace analytique normal (à singularités de quotient

finies, c'est-à-dire localement le quotient d'un ouvert de $\mathbb{C}^k$ par l'action holomorphe d'un groupe fini), et même de variété complexe si $n \geq 3$.

(e) Pour tout diviseur positif m de n, l'identité sur $\mathcal{H}_g$ induit par passage au quotient un morphisme fini $A_g(n)_{\mathbb{C}} \to A_g(m)_{\mathbb{C}}$ qui correspond en termes d'espaces de module, à multiplier la base symplectique de $X[n]$ par n/m pour obtenir une base symplectique de $X[m]$.

Démonstration. Pour le (a) et le (b), c'est une vérification directe découlant principalement de la formule pour les évaluations de ω_τ sur la $\mathbb{Z}$-base naturelle $(e_1, \cdots, e_g, e_1\tau, \cdots, e_g\tau)$ de Λ_τ (c'est bien une base adaptée pour ω_τ).

Pour le (c), l'action est proprement discontinue ([Deb99], Proposition VII.1.3). Ensuite, si $\gamma \in \Gamma_g(n)$ avec $n \geq 3$ est tel que $\gamma \cdot \tau = \tau$, cela signifie que l'automorphisme de A_τ induit par γ fixe à la fois ω_τ et $\alpha_{\tau,n}$, et il est alors égal à l'identité par la proposition VII.3.4 de [Deb99], donc $\gamma = 1$. L'action de $\Gamma_g(n)$ pour $n \geq 3$ est ainsi libre et proprement discontinue, donc librement discontinue.

Pour le (d), c'est l'objet du théorème 8.3.2 de [BL04] (voir les sections 8.1, 8.2 et 8.3 de ce livre pour les détails) pour n quelconque. Dans le cas où $n \geq 3$, on obtient une variété complexe et pas seulement un espace analytique normal grâce au fait que l'action de $\Gamma_g(n)$ est librement discontinue.

Enfin, le (e) se vérifie immédiatement. □

Dans le cas où $n \geq 3$, on peut faire mieux que construire l'espace de modules $A_g(n)_{\mathbb{C}}$, on peut même construire la variété abélienne complexe principalement polarisée universelle, avec structure de niveau n.

Définition-Proposition III.5.4 (Variété abélienne universelle et fibré des formes modulaires)**.** Soit un entier $g \geq 1$.

(a) Le produit semi-direct $\Gamma_g(1) \ltimes \mathbb{Z}^{2g}$ défini par l'action à droite naturelle de $\Gamma_g(1)$ sur $\mathbb{Z}^{2g}$ agit naturellement à gauche par biholomorphismes sur $\mathcal{H}_g \times \mathbb{C}^g$ par la formule

$$(\gamma, (p, q)) \cdot (\tau, z) = \left(\gamma \cdot \tau, (z + p\tau + q) j_\gamma(\tau)^{-1}\right),$$

et cette action est proprement discontinue.

(b) Pour tout entier $n \geq 3$, le quotient $\mathcal{X}_g(n) = \Gamma_g(n) \ltimes \mathbb{Z}^{2g} \backslash \mathcal{H}_g \times \mathbb{C}^g$ pour cette action est une variété complexe munie d'une fibration naturelle $\pi : \mathcal{X}_g(n) \to A_g(n)_{\mathbb{C}}$ telle que pour tout $\tau \in \mathcal{H}_g$, la fibre au-dessus de l'image de τ est la variété abélienne polarisée A_τ.

(c) Pour tout entier $n \geq 3$, le quotient L de $\mathcal{H}_g \times \mathbb{C}$ par l'action de $\Gamma_g(n)$ définie par

$$\gamma \cdot (\tau, z) = (\gamma \cdot \tau, z/\det(j_\gamma(\tau)))$$

est un fibré en droites sur $A_g(n)_{\mathbb{C}}$ appelé *fibré des formes modulaires pour* $\Gamma_g(n)$. Pour tout entier $k \geq 1$, les sections globales de $L^{\otimes k}$ correspondent aux formes modulaires de Siegel de degré g, de poids k et de niveau n (définies par exemple dans [Kli90]).

Pour $n = 1$ ou 2, il existe un exposant k tel que si $\gamma \cdot \tau = \tau$, alors $\det(j_\gamma(\tau))^k = 1$ et on peut donc définir non pas L mais sa puissance k-ième comme un fibré en droites sur $A_g(n)$. On définit alors L comme un $\mathbb{Q}$-fibré en droites (c'est-à-dire comme un fibré en droites formellement multiplié par un rationnel, de sorte qu'un de ses multiples est un vrai fibré en droites).

Démonstration.

(a) Vérifier que c'est une action de groupes est un calcul immédiat, ainsi que le fait qu'elle est proprement discontinue. Remarquons que restreinte à $\mathbb{Z}^{2g}$, c'est une action par translation, et le quotient nous donne alors une fibration au-dessus de $\mathcal{H}_g$ telle que la fibre au-dessus de τ est A_τ par définition, et que restreinte à $\Gamma_g(1)$, sa première coordonnée est l'action de $\Gamma_g(1)$ sur $\mathcal{H}_g$, et la seconde correspond à l'isomorphisme de variétés abéliennes polarisées induit par un γ : plus explicitement, si $\gamma \cdot \tau = \tau'$, l'automorphisme $z \mapsto z j_\gamma(\tau)^{-1}$ de $\mathbb{C}^g$ envoie Λ_τ sur $\Lambda_{\tau'}$ et ω_τ sur $\omega_{\tau'}$, donc induit un isomorphisme de variétés abéliennes principalement polarisées $A_\tau \cong A_{\tau'}$.

(*b*) Comme $n \geq 3$, d'après la définition-proposition III.5.3 (*c*), l'action de $\Gamma_g(n) \ltimes \mathbb{Z}^{2g}$ est sans point fixe (car elle l'est sur la première coordonnée, et les translations sont également sans point fixe). Maintenant, l'application de projection naturelle $\mathcal{H}_g \times \mathbb{C}^g$ passe au quotient en l'application $\pi : \mathcal{X}_g(n) \to A_g(n)_{\mathbb{C}}$, et il reste donc à étudier la fibre au-dessus de chaque point. Pour tout $x \in A_g(n)_{\mathbb{C}}$, il existe $\tau \in \mathcal{H}_g$ tel que x est l'orbite de τ pour $\Gamma_g(n)$, et alors, pour tout $(\tau', z') \in \mathcal{H}_g \times \mathbb{C}^g$, si $\pi(\overline{(\tau', z')}) = x$, il existe un unique $\gamma \in \Gamma_g(n)$ tel que $\gamma \cdot \tau = \tau'$. Ainsi, la fibre au-dessus de x est exactement constituée de tous les $\overline{(\tau, z)}$ avec $z \in \mathbb{C}^g$, mais $\overline{(\tau, z)} = \overline{(\tau, z')}$ si et seulement si z et z' sont égaux à Λ_τ près (car $\gamma \cdot \tau = \tau$ si et seulement si $\gamma = 1$ car l'action de $\Gamma_g(n)$ est libre), ce qui prouve que la fibre au-dessus de $\overline{\tau}$ est bien A_τ.

(*c*) Les arguments pour $n \geq 3$ sont les mêmes que ci-dessus, et les sections globales de $L^{\otimes k}$ se relèvent par construction (comme $\mathcal{H}_g$ est simplement connexe) en des fonctions f holomorphes sur $\mathcal{H}_g$ telles que pour tout $\gamma \in \Gamma_g(n)$ et tout $\tau \in \mathcal{H}_g$,

$$f(\gamma \cdot \tau) = \det(j_\gamma(\tau))^k f(\tau),$$

c'est-à-dire qu'elles s'identifient naturellement aux formes modulaires de Siegel de degré g, de niveau n et de poids k. □

Remarque III.5.2. Ce qui change le (*b*) pour $n = 1$ ou 2 est que $-1 \in \Gamma_g(n)$, et donc le γ tel que $\gamma\tau = \tau'$ n'est pas unique. Ceci conduit à devoir identifier dans chaque fibre (τ, z) et $(\tau, -z)$: ainsi, les fibres au-dessus d'un τ donné sont en général les $A_\tau/[\pm 1]$, et même des quotients par un plus grand groupe (fini) pour certains τ.

Les espaces $A_g(n)_{\mathbb{C}}$ ne sont pas compacts, et il est donc important (ainsi que pour les voir comme des variétés projectives) de les compactifier. Il existe plusieurs façons de le faire, nous allons commencer par la plus naturelle, c'est-à-dire la compactification de Satake. Pour cela, nous allons utiliser un autre modèle complexe de $\mathcal{H}_g$, nous basant sur (et détaillant) la remarque 4.5 de [Nam80].

Définition-Proposition III.5.5 (Autre modèle de $\mathcal{H}_g$)**.**
Soit un entier $g \geq 1$. Sur $\mathbb{C}^{2g}$, on note B la forme $\mathbb{C}$-bilinéaire alternée définie par la matrice J et H la forme hermitienne définie par $H(x, y) := iB(x, \overline{y})$. Pour une matrice $M \in M_{2g,g}(\mathbb{C})$, on va noter $\mathrm{Vect}(M)$ le sous-espace vectoriel de $\mathbb{C}^{2g}$ (vu comme vecteurs colonnes ici) engendré par les colonnes de M.

(*a*) L'application $\tau \mapsto \mathrm{Vect}\begin{pmatrix}\tau \\ 1\end{pmatrix}$ est un isomorphisme de variétés complexes $\Gamma_g(1)$-équivariant entre $\mathcal{H}_g$ et l'ensemble $\mathcal{D}_g \subset \mathrm{Grass}_{g,2g}(\mathbb{C})$ des $\mathbb{C}$-espaces vectoriels W de dimension g de $\mathbb{C}^{2g}$ tels que $B_{|W} \equiv 0$ et $H_{|W}$ est définie positive. On note $\overline{\mathcal{D}_g}$ l'ensemble des $\mathbb{C}$-espaces vectoriels W de dimension g de $\mathbb{C}^{2g}$ tels que $B_{|W} \equiv 0$ et $H_{|W}$ est positive.

(*b*) Pour tout $W \in \overline{\mathcal{D}_g}$, l'espace vectoriel $W \cap \overline{W}$ est égal à $V \otimes_{\mathbb{R}} \mathbb{C}$ pour un certain sous-$\mathbb{R}$-espace vectoriel V de $\mathbb{R}^{2g}$, et ce V est totalement isotrope pour B et H. Pour tout sous-espace vectoriel V de $\mathbb{R}^{2g}$ isotrope pour B (et donc H), on appelle *composante de bord associée à* V l'ensemble

$$\mathcal{C}(V) := \{W \in \overline{\mathcal{D}_g} \,|\, W \cap \overline{W} = V \otimes_{\mathbb{R}} \mathbb{C}\},$$

et on dit que celle-ci est de degré g' si V est de dimension g'. De plus, pour tout $\gamma \in \Gamma_g(1)$, $\gamma \cdot \mathcal{C}(V) = \mathcal{C}(\gamma \cdot V)$. En particulier, pour tout entier $g' \in \{0, \cdots, g\}$, on note $\mathcal{C}_{g'} := \mathcal{C}(V_{g'})$ où $V_{g'} = \mathrm{Vect}\begin{pmatrix}0 \\ 1_{g'} \\ 0_g\end{pmatrix}$ (par exemple, $\mathcal{C}_0 = \mathcal{D}_g$). On a alors

$$\mathcal{C}_{g'} = \left\{ \mathrm{Vect}\begin{pmatrix}\tau' & 0 \\ 0 & 1_{g'} \\ 1_{g-g'} & 0 \\ 0 & 0\end{pmatrix}, \tau' \in \mathcal{H}_{g-g'} \right\},$$

et on note $\phi_{g'} : \mathcal{H}_{g-g'} \to \mathcal{C}_{g'}$ la bijection ainsi induite.

(*c*) Une composante de bord $\mathcal{C}(V)$ est *rationnelle* si V est défini sur $\mathbb{Q}$, ou de manière équivalente si $V = \gamma \cdot V_{g'}$ avec $\gamma \in \Gamma_g(1)$, pour g' le degré de V. On note $\mathcal{D}_g^c \subset \overline{\mathcal{D}_g}$ l'union de $\mathcal{D}_g$ et de toutes les composantes de bord rationnelles de $\overline{\mathcal{D}_g}$.

(*d*) L'action de $\Gamma_g(1)$ sur $\mathcal{H}_g$ s'étend naturellement en une action de $\Gamma_g(1)$ sur $\mathcal{D}_g^c$.

(*e*) Pour tout $g' \in \{0, \cdots, g\}$, le stabilisateur de la composante de bord $\mathcal{C}_{g'}$ est constitué des matrices $\gamma \in \Gamma_g(1)$ de la forme

$$\begin{pmatrix} A'' & 0 & B'' & * \\ * & u & * & * \\ C'' & 0 & D'' & * \\ 0 & 0 & 0 & {}^t u^{-1} \end{pmatrix} \quad \text{avec} \quad \begin{pmatrix} A'' & B'' \\ C'' & D'' \end{pmatrix} \in \Gamma_{g-g'}(1) \text{ et } u \in \mathrm{GL}_{g'}(\mathbb{Z}).$$

Démonstration. Tout d'abord, pour tout $W \in \mathcal{D}_g^c$, le $\mathbb{C}$-espace vectoriel $W \cap \overline{W}$ est stable par conjugaison complexe et il existe donc bien un unique sous-espace vectoriel réel V de $\mathbb{R}^{2g}$ tel que $W \cap \overline{W} = V \otimes_{\mathbb{R}} \mathbb{C}$. Ensuite, les éléments de $W \cap \overline{W}$ sont exactement les $x \in \mathbb{C}^{2g}$ orthogonaux à W à la fois relativement à B et à H : en effet, W est de dimension g et B bilinéaire alternée non dégénérée, donc c'est un sous-espace totalement isotrope maximal pour B et son orthogonal pour H est exactement $\overline{W}$. Comme H est positive sur W, l'intersection $W \cap \overline{W}$ est constituée des éléments de W isotropes pour H.

Ensuite, pour $g' \in \{0, \cdots, g\}$, notons $g'' = g - g'$ et supposons que $W \in \mathcal{C}_{g'}$. Soit un vecteur $v \in \mathbb{C}^{2g}$ dont les coordonnées d'indice $g''+1, \cdots, g''+g$ sont nulles. Remarquons qu'il ne peut alors pas appartenir à $V_{g'}$ à moins d'être nul par définition de celui-ci, et est isotrope pour B et H. S'il appartient à W, alors comme il est isotrope pour H, il appartient à $W \cap \overline{W}$ et donc $v = 0$. Ceci prouve que la projection de W dans $\mathbb{C}^g$ qui à un vecteur associe ses coordonnées $g''+1, \cdots, g''+g$ est surjective (car injective), et on peut donc écrire $W = W_\tau = \mathrm{Vect}\, M_\tau$, avec

$$M_\tau = \begin{pmatrix} \tau'' & \tau_2 \\ 0 & 1_{g'} \\ 1_{g-g'} & 0 \\ -\tau_3 & \tau' \end{pmatrix}, \tau = \begin{pmatrix} \tau'' & \tau_2 \\ \tau_3 & \tau' \end{pmatrix}, \tau' \in M_{g'}(\mathbb{C}), \tau'' \in M_{g''}(\mathbb{C}).$$

Maintenant, par calcul matriciel on voit que pour tous vecteurs colonnes $X, Y \in \mathbb{C}^g$:

$$B\left(M_\tau X, M_\tau Y\right) = {}^t X \begin{pmatrix} \tau'' - {}^t\tau'' & {}^t\tau_3 - \tau_2 \\ {}^t\tau_2 - \tau_3 & \tau' - {}^t\tau' \end{pmatrix} Y \text{ et } H\left(M_\tau X, M_\tau Y\right) = {}^t X \begin{pmatrix} \overline{\tau''} - {}^t\tau'' & {}^t\tau_3 - \overline{\tau_2} \\ {}^t\tau_2 - \overline{\tau_3} & \overline{\tau'} - {}^t\tau' \end{pmatrix} Y.$$

Ceci prouve que $W_\tau \in \mathcal{D}_g^c$ si et seulement si τ est symétrique et $\mathrm{Im}\,\tau$ positive, et dans ce cas

$$H\left(M_\tau X, M_\tau Y\right) = 2 {}^t X \,\mathrm{Im}\,\tau \overline{Y}.$$

Plus précisément, on en déduit que l'espace vectoriel réel V_τ associé à W_τ est l'ensemble des produits matriciels $M_\tau X$ avec $X \in \mathbb{R}^g$ tel que $(\mathrm{Im}\,\tau)X = 0$. Ainsi, on sait que $V_\tau = V_{g'}$ si et seulement si $\mathrm{Ker}(\mathrm{Im}\,\tau) = \mathrm{Vect}(e_{g-g'+1}, \cdots, e_g)$ et que la multiplication de cet espace vectoriel par $\mathrm{Re}\,\tau$ est nulle, ce qui équivaut à dire que τ est de la forme

$$\tau = \begin{pmatrix} \tau'' & 0 \\ 0 & 0 \end{pmatrix}, \tau'' \in \mathcal{H}_{g''},$$

ce qui prouve le (a) et le (b). La nature $\Gamma_g(1)$-équivariante de l'application du (a) est naturelle car pour tout $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_g(1)$,

$$\mathrm{Vect}\begin{pmatrix} \gamma \cdot \tau \\ 1_g \end{pmatrix} = \mathrm{Vect}\begin{pmatrix} A\tau + B \\ C\tau + D \end{pmatrix} = \gamma \cdot \mathrm{Vect}\begin{pmatrix} \tau \\ 1_g \end{pmatrix},$$

et l'action naturelle de $\Gamma_g(1)$ sur $\mathbb{C}^{2g}$ préserve B et H donc $\Gamma_g(1)$ préserve $\mathcal{D}_g^c$ dans $\mathrm{Grass}_{g,2g}(C)$.

Pour le (c), soit $(e_1, \cdots, e_{2g})$ la base canonique de $\mathbb{Z}^{2g}$. Soit V un sous-espace vectoriel réel de $\mathbb{R}^{2g}$ défini sur $\mathbb{Q}$ et totalement isotrope pour B. Il existe donc une $\mathbb{R}$-base $(f_1, \cdots, f_{g'})$ de V dont chacun des éléments appartient à $\mathbb{Q}^{2g}$, et orthogonaux deux à deux pour B. On peut même, quitte à multiplier soigneusement ces éléments, supposer qu'ils appartiennent à $\mathbb{Z}^{2g}$ et sont complétables en une $\mathbb{Z}$-base de $\mathbb{Z}^{2g}$, mais ils sont alors complétables en une base symplectique de $\mathbb{Z}^{2g}$. Comme $\Gamma_g(1)$ agit transitivement sur les bases symplectiques de $\mathbb{Z}^{2g}$ par construction, il existe γ tel que $(f_1, \cdots, f_{g'}) = \gamma \cdot (e_{g-g'+1}, \cdots, e_g)$ et donc

$$V = \mathrm{Vect}(f_1, \cdots, f_{g'}) = \gamma \cdot \mathrm{Vect}(e_{g-g'+1}, \cdots, e_g) = \gamma V_{g'}.$$

Réciproquement, un tel espace vectoriel est bien défini sur $\mathbb{Q}$.

Enfin, pour le (e), il suffit de décomposer une matrice γ en seize blocs comme pour la forme de l'énoncé et de calculer le produit de γ par $\begin{pmatrix} 0 \\ 1_{g'} \\ 0_g \end{pmatrix}$, puis de vérifier quand il appartient à $V_{g'}$. On obtient immédiatement l'annulation des trois blocs de la seconde colonne sauf le second qui doit être inversible car γ doit induire une bijection sur le $\mathbb{Z}$-module sous-jacent à $V_{g'}$. Le reste s'en déduit par symplecticité (pour plus de détails, voir la proposition 4.8 de [Nam80]). □

Cette série de définitions va nous permettre de construire (suivant la définition 5.7 de [Nam80]) une topologie sur $\mathcal{D}_g^c$, dont la topologie quotient nous donnera la compactification de Satake.

Définition III.5.6 (Topologie cylindrique sur $\mathcal{D}_g^c$)**.**

Soit $g \geq 1$ un entier.

(a) Pour tout entier $0 \leq g' \leq g$, tout ouvert U de $\mathcal{H}_{g'}$ pour la topologie usuelle et tout réel $\lambda > 0$, on définit la partie $\mathcal{U}(U, \lambda)$ de $\mathcal{D}_g^c$ par :

$$\mathcal{U}(U, \lambda) := \bigcup_{0 \leq g_1 \leq g'} \mathcal{U}(U, \lambda, g_1)$$

où pour chaque $g_1 \leq g'$, l'ensemble $\mathcal{U}(U, \lambda, g_1) \subset \mathcal{C}_{g_1}$ est l'image par ϕ_{g_1} des $\tau \in \mathcal{H}_{g-g_1}$ qui sont de la forme $\tau = \begin{pmatrix} \tau_1 & \tau'' \\ {}^t\tau'' & \tau_2 \end{pmatrix}$, avec

$$\tau_1 \in U \subset \mathcal{H}_{g-g'} \text{ et } \operatorname{Im}\tau_2 - {}^t\operatorname{Im}\tau''(\operatorname{Im}\tau_2)^{-1}\operatorname{Im}\tau'' > \lambda\,\mathrm{Id}_{g'-g_1},$$

où le symbole $>$ signifie que la différence des deux termes est une matrice (symétrique réelle) définie positive. La *topologie cylindrique sur* $\mathcal{D}_g^c$ est la topologie la moins fine contenant comme ouverts tous les $\mathcal{U}(U, \lambda)$ et telle que les $\gamma \in \Gamma_g(1)$ agissent par homéomorphismes.

(b) Pour tout entier $n \geq 1$, la *compactification de Satake de* $A_g(n)_{\mathbb{C}}$, notée $A_g(n)^S_{\mathbb{C}}$, est l'espace $\Gamma_g(n)\backslash\mathcal{D}_g^c$ muni de la topologie quotient de la topologie cylindrique.

Pour mieux comprendre les définitions ci-dessus, donnons quelques propriétés de base de cette topologie et cette compactification.

Proposition III.5.7.

Soit $g \geq 1$ un entier.

(a) La topologie cylindrique induit sur chaque composante de bord $\mathcal{C}_{g'}$ la topologie induite par $\mathcal{H}_{g-g'}$ et $\phi_{g'}$, c'est-à-dire que $\phi_{g'}$ est un homéomorphisme de $\mathcal{H}_{g-g'}$ vers $\mathcal{C}_{g'}$, et il en est de même pour les autres composantes de bord.

(b) Une suite $\tau_n = \begin{pmatrix} \tau_n^{(1)} & \tau_n'' \\ {}^t\tau_n'' & \tau_n^{(2)} \end{pmatrix}$ d'éléments de $\mathcal{H}_g$ avec $\tau_n^{(1)} \in \mathcal{H}_{g-g'}$ pour un certain $g' \leq g$ converge vers $\phi_{g'}(\tau') \in \mathcal{C}_{g'}$ pour un certain $\tau' \in \mathcal{H}_{g-g'}$ dans la topologie cylindrique si et seulement si

$$\tau_n^{(1)} \to \tau' \text{ et } \operatorname{Im}\tau_n^{(2)} - {}^t\operatorname{Im}\tau_n''(\operatorname{Im}\tau_n^{(2)})^{-1}\operatorname{Im}\tau_n'' \to +\infty.$$

c'est-à-dire que pour tout $\lambda > 0$, à partir d'un certain rang n, on a

$$\operatorname{Im}\tau_n^{(2)} - {}^t\operatorname{Im}\tau_n''(\operatorname{Im}\tau_n^{(2)})^{-1}\operatorname{Im}\tau_n'' > \lambda\,\mathrm{Id}_{g'g'}\,.$$

(*c*) *En tant qu'ensemble, on a*

$$\Gamma_g(1)\backslash\mathcal{D}_g^c = A_g(1)_{\mathbb{C}} \sqcup \cdots \sqcup A_1(1)_{\mathbb{C}} \sqcup A_0(1)_{\mathbb{C}}.$$

où $A_0(1)_{\mathbb{C}}$ est réduit à un point.

(*d*) *Pour toute composante de bord rationnelle $\mathcal{C}(V)$, l'adhérence de $\mathcal{C}(V)$ dans $\mathcal{D}_g^c$ pour la topologie cylindrique est exactement l'union des composantes de bord $\mathcal{C}(V')$ avec V' contenant V.*

(*e*) *Pour $g = 1$, l'ensemble $\mathcal{D}_1^c$ s'identifie canoniquement à $\mathcal{H}_1 \cup \mathbb{P}^1(\mathbb{Q})$, et la topologie cylindrique est la topologie dite hyperbolique sur cet espace. Pour tout $n \geq 1$, le quotient $\Gamma_1(n)\backslash\mathcal{H}_1$ s'identifie alors à la courbe modulaire (non compacte) $Y(n)$ paramétrant les classes d'isomorphisme de courbes elliptiques munies d'une structure complète de niveau n, et sa compactification de Satake s'identifie à la courbe modulaire $X(n)$ (section I.5 de [DS05], par exemple).*

Démonstration. (*a*) Par définition, l' intersection d'un $\mathcal{U}(U,\lambda)$ avec $\mathcal{C}_{g'}$ est $\phi_{g'}(U)$, d'où le résultat (il faut aussi voir avec la preuve du (*c*) que chaque $\Gamma_g(1)$ fixant une composante de bord de degré g' donnée agit par homéomorphismes sur cette composante de bord lorsqu'elle est munie de la topologie induite par $\mathcal{H}_{g-g'}$).

(*b*) C'est la Scholie 5.9 p. 38 de [Nam80]. C'est en fait l'idée intuitive derrière la construction des composantes de bord, permettant de lire assez naturellement l'adhérence d'une partie de $\mathcal{H}_g$ dans $\mathcal{D}_g^c$.

(*c*) Par définition des composantes rationnelles, les composantes rationnelles de même degré forment une seule orbite par l'action de $\Gamma_g(1)$, d'où la stratification. Il reste à vérifier que pour tout $g' \in \{0, \cdots, g\}$, l'action du stabilisateur de $\mathcal{C}_{g'}$ se traduit en l'action de $\Gamma_{g'}(1)$ sur $\mathcal{H}_{g-g'}$ via $\phi_{g'}$. D'après les (*b*) et (*e*) de la définition-proposition III.5.5, un élément W de $\mathcal{C}_{g'}$ s'écrit sous la forme $\phi_{g'}(\tau'')$ avec $\tau'' \in \mathcal{H}_{g-g'}$ et pour γ ayant la forme énoncée, on vérifie directement que

$$\gamma \cdot \phi_{g'}(\tau'') = \phi_{g'}((A''\tau + B'')(C''\tau'' + D'')^{-1}).$$

et donc $\phi_{g'}(\tau'')$ et $\phi_{g'}(\tau_2'')$ sont dans la même orbite par $\Gamma_g(1)$ si et seulement si τ'' et τ_2'' le sont par $\Gamma_{g''}(1)$, ce qui conclut la preuve.

(*d*) D'après le (*b*), on voit que l'adhérence de $\mathcal{H}_g = \mathcal{C}_0$ dans $\mathcal{D}_g^c$ contient tous les $\mathcal{C}_{g'}, g' \leq g$ donc toutes les composantes de bord (rationnelles) par définition de celles-ci comme $\Gamma_g(1)$ agit par homéomorphismes. Ainsi, l'espace $\mathcal{H}_g$ est dense dans $\mathcal{C}_0$. Par le même raisonnement appliqué au degré supérieur, comme pour tout $g' \leq g$, l'union des composantes de bord $\mathcal{C}(V')$ avec V' contenant $V_{g'}$ est homéomorphe à $\mathcal{D}_{g-g'}^c$ avec les topologies cylindriques, $\mathcal{C}_{g'}$ est dense dans cette union, et transporter cette propriété par action de $\Gamma_g(1)$ (qui agit transitivement sur les composantes de bord de même degré) donne le (*d*).

(*e*) Il suffit de voir la définition de la topologie cylindrique dans le cas $g = 1$. □

Les propriétés importantes de cette compactification sont les suivantes.

Proposition III.5.8.

Pour tous entiers $g \geq 1$ et $n \geq 1$:

(*a*) *Le groupe $\Gamma_g(n)$ agit proprement discontinument sur $\mathcal{D}_g^c$.*

(*b*) *La compactification de Satake $A_g(n)_{\mathbb{C}}^S$ est un espace topologique compact contenant $A_g(n)_{\mathbb{C}}$ comme ouvert dense, et son complémentaire est de codimension g.*

(*c*) *Cette compactification admet une structure naturelle d'espace analytique normal prolongeant celle de $A_g(n)_{\mathbb{C}}$, et c'est une variété algébrique projective. Plus précisément, pour $M_g(n)$ la $\mathbb{C}$-algèbre graduée (par le poids) des formes modulaires de Siegel de degré g et de niveau n (qui est normale de type fini), les espaces analytiques normaux $A_g(n)_{\mathbb{C}}^S$ et $\operatorname{Proj}_{\mathbb{C}} M_g(n)$ sont isomorphes. De manière plus concrète, cela signifie que la compactification de Satake $A_g(n)_{\mathbb{C}}^S$ est exactement*

l'adhérence dans un certain espace projectif de l'image de $A_g(n)_{\mathbb{C}}$ par l'évaluation en une base de formes modulaires de degré g, de niveau n et de même poids k (assez grand).

(d) Le $\mathbb{Q}$-fibré L des formes modulaires sur $A_g(n)_{\mathbb{C}}$ s'étend canoniquement à $A_g(n)^S_{\mathbb{C}}$ et cette extension est un $\mathbb{Q}$-fibré ample sur la compactification de Satake.

Démonstration. Les assertions (a) et (b) sont contenues dans le théorème 5.10 de [Nam80] (au moins pour la compacité, c'est une conséquence rapide de la proposition III.5.7 (a) ci-dessus et du domaine fondamental de $\Gamma_g(1)\backslash\mathcal{H}_g$ qu'on peut trouver dans l'Exemple 5.12 [Nam80]). Pour la codimension du complémentaire, il suffit de remarquer qu'en tant que variété complexe, chaque composante de bord de degré g' est isomorphe à $A_{g-g'}(n)_{\mathbb{C}}$ donc de dimension $(g-g')(g-g'+1)/2$, ainsi le bord est de dimension $(g-1)g/2$, alors que $A_g(n)_{\mathbb{C}}$ est de dimension $g(g+1)/2$. Pour le (c), la structure naturelle d'espace analytique normal a été établie par [CS57], et l'isomorphisme d'espaces analytiques normaux par le théorème 5 de [Bai58] dans le cas $n=1$ et le « théorème fondamental » de [Car57] dans le cas général (essentiellement à l'aide de la proposition 12 de cet article et du résultat de [Bai58]). Pour le (d), la nature même de l'espace $\mathrm{Proj}_{\mathbb{C}} M_g(n)$ et de l'isomorphisme (obtenu par évaluation des formes modulaires d'un poids suffisamment grand) montre que le $\mathbb{Q}$-fibré L s'étend canoniquement à $\mathrm{Proj}_{\mathbb{C}} M_g(n)$, donc à $A_g(n)^S_{\mathbb{C}}$, et a une certaine puissance très ample, il est donc ample. □

Remarque III.5.3. Nous verrons (définition-proposition III.5.22 (b))que beaucoup de formes modulaires de degré g peuvent être construites comme fonctions thêta, et pour certaines structure de niveau, ces fonctions thêta engendrent bien (à partir d'un certain poids) l'algèbre graduée des formes modulaires avec cette structure : ainsi, comme l'a démontré pour la première fois Igusa (voir [Igu72], section V.4 et théorème 4 de cette section ou encore [Deb99], Théorème 4.1), pour certaines structures de niveau, il suffit d'évaluer en tout point par les fonctions thêta (ou plutôt « thêta-constantes ») adaptées pour obtenir un plongement quasi-projectif, et de là la compactification de Satake. En général, les fonctions thêta qu'on peut construire ne suffisent cependant pas et il faut utiliser toutes les formes modulaires possibles.

Voyons maintenant quelles sont exactement les différentes composantes de bord dans $A_g(n)^S_{\mathbb{C}}$, et leurs adhérences.

Proposition III.5.9.

Pour tous entiers $g \geq 1$ et $n \geq 1$:

(a) Deux composantes de bord $\mathcal{C}(V)$ et $\mathcal{C}(V')$ de $\mathcal{D}^c_g$ sont identifiées dans $A_g(n)^S_{\mathbb{C}}$ si et seulement s'il existe $\gamma \in \Gamma_g(n)$ tel que $\gamma \cdot V = V'$. On appellera composante de bord de $A_g(n)^S_{\mathbb{C}}$ la projection dans $A_g(n)^S_{\mathbb{C}}$ d'une composante de bord de $\mathcal{D}^c_g$, et si elle est de degré g', elle est isomorphe en tant qu'espace analytique à $A_{g-g'}(n)_{\mathbb{C}}$.

(b) Soit $g' \in \{0, \cdots, g\}$. Alors, l'application qui à une matrice M de $M_{2g,g'}(\mathbb{Z})$ complétable en une matrice symplectique de taille $2g$ associe la composante de bord $\mathcal{C}(\mathrm{Vect}\, M)$ induit après passage au quotient modulo n, une surjection de l'ensemble des matrices $M \in M_{2g,g'}(\mathbb{Z}/n\mathbb{Z})$ complétables en des matrices symplectiques de $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$ vers l'ensemble des composantes de bord de degré g' de $A_g(n)^S_{\mathbb{C}}$. On note, pour tout $M \in M_{2g,g'}(\mathbb{Z}/n\mathbb{Z})$ ayant ces propriétés, $\mathcal{C}(M)$ la composante correspondante, et alors $\mathcal{C}(M) = \mathcal{C}(M')$ si et seulement s'il existe une matrice u de $\mathrm{GL}_{g'}(\mathbb{Z})$ (vue modulo n) telle que $M = M'u$. En particulier, les composantes de bord de degré 1 de $A_g(n)^S_{\mathbb{C}}$ correspondent aux vecteurs colonnes primitifs de $(\mathbb{Z}/n\mathbb{Z})^{2g}$, à ± 1 près, il y en a donc $n^{2g}/2\prod_p(1-p^{-2g})$ (où p parcourt les nombres premiers divisant n) si $n \geq 3$.

(c) Via les correspondances du (b), l'adhérence de la composante de bord $\mathcal{C}(M')$ de degré g' est constituée de l'union des $\mathcal{C}(M_1)$ où M_1 est une matrice complétable en une matrice symplectique de $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$, dont les g' premières colonnes forment M'.

Démonstration. (a) C'est dû au fait que l'action de $\Gamma_g(1)$ se fait composante de bord par composante de bord et est transitive entre celles de même degré, et grâce à la formule donnée par la preuve de la proposition III.5.7 (c).

(b) D'après la proposition III.5.7 (b) et sa preuve, on sait que $\Gamma_g(1)$ agit transitivement sur toutes les composantes de bord de même degré. Étant donné deux telles composantes, on veut donc exprimer quand elles appartiennent à la même orbite non pas dans $\Gamma_g(1)$ mais dans $\Gamma_g(n)$.

À toute composante de bord rationnelle $\mathcal{C}(V)$ de $\mathcal{D}_g^c$, on peut associer une matrice M de $M_{2g,g'}(\mathbb{Z})$ dont les colonnes engendrent le réseau $V \cap \mathbb{Z}^{2g}$, de dimension g' et isotrope pour B. On note $\mathcal{M}_{g,g'}$ l'ensemble de ces matrices, et deux matrices M et M' de $\mathcal{M}_{g,g'}$ correspondent à la même composante de bord dans $\mathcal{D}_g^c$ si et seulement si $M = M'u$ avec $u \in \mathrm{GL}_{g'}(\mathbb{Z})$. Maintenant, deux composantes de bord $\mathcal{C}(\mathrm{Vect}\, M)$ et $\mathcal{C}(\mathrm{Vect}\, M')$ s'identifient dans $A_g(n)^S_{\mathbb{C}}$ si et seulement s'il existe $\gamma \in \Gamma_g(n)$ et $u \in \mathrm{GL}_{g'}(\mathbb{Z})$ tels que $M' = \gamma M u$, ainsi l'ensemble des composantes de bord de degré g' de la compactification de Satake s'identifie naturellement au double quotient

$$\Gamma_g(n)\backslash \mathcal{M}_{g,g'} / \mathrm{GL}_{g'}(\mathbb{Z}),$$

sur lequel agit canoniquement le groupe $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$. Ensuite, le groupe $\Gamma_g(1)$ se surjecte dans $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$ pour la projection canonique : il suffit de montrer que ce dernier est engendré par les réductions modulo n de matrices élémentaires qui sont elles-mêmes symplectiques dans $\mathbb{Z}$ (ces matrices peuvent être trouvées dans l'Annexe de [Mum87], p.202 et la preuve qu'elles engendrent $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$ et $\mathrm{Sp}_{2g}(\mathbb{Z})$ sont très similaires). Deux matrices M et M' de $\mathcal{M}_{g,g'}$ sont donc dans la même orbite pour $\Gamma_g(n)$ si et seulement si elles sont congrues modulo n, donc l'ensemble des composantes de bord s'identifie naturellement à

$$\overline{\mathcal{M}_{g,g'}} / \mathrm{GL}_{g'}(\mathbb{Z}),$$

où $\overline{\mathcal{M}_{g,g'}}$ est la projection modulo n de $\mathcal{M}_{g,g'}$, mais c'est aussi l'ensemble des matrices de $M_{2g,g'}(\mathbb{Z}/n\mathbb{Z})$ complétables en une matrice de $\mathrm{Sp}_{2g}(\mathbb{Z}/n\mathbb{Z})$. Dans le cas où $g' = 1$, l'ensemble $\mathcal{M}_{g,g'}$ est simplement constitué des vecteurs entiers primitifs de $(\mathbb{Z}/n\mathbb{Z})^{2g}$, d'où le résultat.

(c) C'est une conséquence directe de la proposition III.5.7 (d) et du (a).

□

Cette compactification a le défaut, pour $g > 1$, de ne jamais être lisse (même si $n \geq 3$), avec un bord de dimension 1. Pour pallier ce problème, Igusa a conçu dans le cas $g = 2$ un éclatement de cette compactification le long du bord, appelé la compactification d'Igusa [Igu67a], et les idées employées ont donné naissance à une méthode générale de compactification, dite toroïdale, développée par plusieurs auteurs notamment Mumford et Tai, et dont on peut trouver une idée plus générale dans [AMRT10].

Un point important est que la compactification toroïdale de $A_g(n)_{\mathbb{C}}$ n'est en général pas unique (et chaque compactification a différentes propriétés), mais pour ce qui suivra nous aurons choisi par défaut $g = 2$ ou 3, et la compactification toroïdale associée à la deuxième décomposition de Voronoï (pour $g = 2$ qui constitue l'essentiel de notre propos, c'est la même que celle d'Igusa, car alors la décomposition en cônes centraux est la même que la deuxième décomposition de Voronoï).

Nous n'utiliserons pas la compactification toroïdale à proprement parler dans cette thèse, mais nous en consignons ici succinctement les propriétés importantes en vue d'applications ultérieures de la méthode de Runge.

La notion de compactification toroïdale passe par les compactifications partielles le long des composantes de bord, que nous expliquons ci-dessous pour $\mathcal{C}_1$ (la compactification partielle le long des composantes de bord de degré plus grand étant plus complexe à décrire).

D'après la définition-proposition III.5.5 (e), le stabilisateur noté P_g de $\mathcal{C}_1$ dans $\Gamma_g(1)$ est constitué des matrices γ de la forme

$$\begin{pmatrix} A'' & 0 & B'' & * \\ * & \pm 1 & * & * \\ C'' & 0 & D'' & * \\ 0 & 0 & 0 & \pm 1 \end{pmatrix} \quad \text{avec} \quad \begin{pmatrix} A'' & B'' \\ C'' & D'' \end{pmatrix} \in \Gamma_{g-g'}(1).$$

En fait, une matrice $\gamma \in \Gamma_g(1)$ appartient à P_g si et seulement si elle stabilise tout $\mathcal{U}(\mathcal{H}_{g-1}, \lambda)$ dès que le réel λ est assez grand : rappelons que ceux-ci sont des voisinages fondamentaux de $\mathcal{C}_1$

pour la topologie cylindrique (définition III.5.6), et que l'action de $\Gamma_g(1)$ est proprement discontinue sur $\mathcal{D}_g^c$ (proposition III.5.8 (a)).

Il y alors une application naturelle

$$\begin{array}{ccc} \mathcal{U}(\mathcal{H}_{g-1},\lambda) & \longrightarrow & \mathcal{H}_{g-1}\times\mathbb{C}^{g-1}\times\mathbb{C}^* \\ \phi_1\begin{pmatrix}\tau_1 & z\\ {}^t z & \tau_2\end{pmatrix} & \longmapsto & (\tau_1, z, e^{2i\pi\tau_2}) \end{array},$$

qui est équivariante pour le groupe P_g lorsque ce dernier agit sur $\mathcal{H}_{g-1}\times\mathbb{C}^{g-1}\times\mathbb{C}^*$, pour les matrices particulières suivantes (le tout formant bien une action cohérente), par

$$\begin{pmatrix} A & 0 & B & 0\\ 0 & 1 & 0 & 0\\ C & 0 & D & 0\\ 0 & 0 & 0 & 1\end{pmatrix}(\tau_1,z,q) = (\gamma\cdot\tau_1, j_\gamma(\tau_1)^{-1}z, qe^{-2i\pi {}^t j_\gamma(\tau)^{-1}Cz}),\ \text{où } \gamma = \begin{pmatrix}A & B\\ C & D\end{pmatrix}\in\Gamma_{g-1}(1).$$

et

$$\begin{pmatrix} 1_{g-1} & 0 & 0 & \ell\\ m & 1 & \ell & 0\\ 0 & 0 & 1_{g-1} & -m\\ 0 & 0 & 0 & 1\end{pmatrix}(\tau_1,z,q) = (\tau_1, z+\tau_1 m+\ell, qe^{2i\pi(m\tau_1 m+2mz)}).$$

(pour plus de détails, voir la section 3.11 de [BvdGHZ08]). Cette action s'étend naturellement en une action sur $\mathcal{H}_{g-1}\times\mathbb{C}^{g-1}\times\mathbb{C}$, et la compactification partielle de $A_g(n)^S_{\mathbb{C}}$ le long de $\mathcal{C}_1$ est alors le quotient de $\mathcal{H}_{g-1}\times\mathbb{C}^g\times\mathbb{C}$ par le groupe $P_{g,1}\cap\Gamma_g(n)$ pour celle-ci. En particulier, $\mathcal{H}_{g-1}\times\mathbb{C}^g\times\{0\}$ est stabilisé par ce groupe, et on remarque que l'action se voit alors exactement comme l'action de $\Gamma_{g-1}(n)\ltimes\mathbb{Z}^{2g-2}$ sur $\mathcal{H}_{g-1}\times\mathbb{C}^{g-1}$ décrite dans la définition-proposition III.5.4. La compactification partielle le long de $\mathcal{C}_1$ fait donc apparaître, d'après celle-ci, la variété abélienne universelle de dimension $g-1$ avec structure de niveau n.

Plus précisément, la forme de la compactification toroïdale que nous allons manipuler est la suivante.

Définition-Proposition III.5.10 (Compactification toroïdale)**.** Soit $g\in\{2,3\}$ et $n\geq 1$ des entiers.

(a) On appelle *compactification toroïdale de* $A_g(n)_{\mathbb{C}}$, notée $A_g^*(n)_{\mathbb{C}}$, l'espace analytique normal obtenu par la méthode de compactification toroïdale associée à la seconde décomposition polyédrale de Voronoï (voir le chapitre 9 de [Nam80] pour des détails).

(b) La variété normale $A_g(n)_{\mathbb{C}}$ est un ouvert dense de $A_g^*(n)_{\mathbb{C}}$ dont le complémentaire est de codimension 1. On appelle *diviseur de bord*, noté D, la somme des composantes irréductibles de ce complémentaire, en tant que diviseur de Weil.

(c) L'espace analytique $A_g(n)^*_{\mathbb{C}}$ est une variété projective complexe, munie d'un morphisme surjectif π vers $A_g(n)^S_{\mathbb{C}}$ égal à l'identité sur $A_g(n)_{\mathbb{C}}$, et pour m divisant n, d'un morphisme fini $A_g(n)^*_{\mathbb{C}}\to A_g(m)^*_{\mathbb{C}}$ prolongeant le morphisme naturel $A_g(n)_{\mathbb{C}}\to A_g(m)_{\mathbb{C}}$ de la définition-proposition III.5.3 (e).

(d) Pour $n\geq 3$, cette variété est lisse, et pour toute composante de bord $\mathcal{C}$ de degré 1 de $A_g(n)^S_{\mathbb{C}}$, on a un diagramme commutatif d'espaces analytiques normaux

$$\begin{array}{ccc} \pi^{-1}(\mathcal{C}) & \xrightarrow{\cong} & \mathcal{X}_{g-1}(n)\\ {\scriptstyle\pi}\downarrow & & \downarrow\\ \mathcal{C} & \xrightarrow{\cong} & A_{g-1}(n)_{\mathbb{C}}\end{array}$$

où les flèches non nommées sont canoniques (définition-proposition III.5.4 (b) pour la construction de $\mathcal{X}_{g-1}(n)$).

Pour $n=1$ ou 2, la variété $A_g^*(n)_{\mathbb{C}}$ est normale à singularités de quotient finies.

(*e*) Le $\mathbb{Q}$-fibré en droites L sur $A_g(n)_{\mathbb{C}}$ s'étend canoniquement sur $A_g^*(n)_{\mathbb{C}}$ comme le pullback par $\pi_n : A_g^*(n)_{\mathbb{C}} \to A_g(n)_{\mathbb{C}}^S$ du $\mathbb{Q}$-fibré L sur $A_g(n)_{\mathbb{C}}^S$ (il prolonge bien le fibré des formes modulaires sur $A_g(n)_{\mathbb{C}}$).

Démonstration. (*b*) La normalité de $A_g^*(n)_{\mathbb{C}}$ découle du processus de compactification toroïdale (théorème 7.12 de [Nam80] par exemple), où à chaque étape de compactification partielle on a affaire à des espaces analytiques normaux. Pour la dimension du complémentaire, cela découle par exemple de la dimension de la compactification partielle le long de $\mathcal{C}_1$: comme celle-ci est la variété abélienne universelle de dimension $g-1$ et de niveau n comme on l'a vu ci-dessus, sa dimension est $g(g-1)/2 + (g-1)$ donc $g(g+1)/2 - 1$.

(*c*) Le fait que $A_g^*(n)_{\mathbb{C}}$ est projective découle de notre choix de compactification toroïdale, qui coïncide avec la compactification d'Igusa pour $g = 2$ et 3, donc la projectivité vient de la construction de celle-ci ([Igu67a], ou théorème 1.2.18 de [Ale02]). Le reste est vrai pour toutes les compactifications toroïdales, et dû au fait qu'en un sens la compactification de Satake est « minimale » (ce qui découle de la construction même et se voit dans le cas de $\mathcal{C}_1$ ci-dessus (voir [Nam80], chapitre 2 pour les résultats précis et les chapitres 6 et 7 pour leurs preuves).

(*d*) Seule la lissité est à prouver, et la seconde décomposition de Voronoï étant régulière pour $g \leq 4$, d'après la proposition 7.19 de [Nam80], il suffit de prouver que $\Gamma_g(n)$ est « net » pour $n \geq 3$, c'est-à-dire que le groupe engendré par les valeurs propres d'un élément de $\Gamma_g(n)$ ne contient jamais de racine de l'unité différente de 1. Pour cela, soit $M \in \Gamma_g(n)$ de polynôme caractéristique P, et K le corps de nombres engendré par ses valeurs propres. Comme M est symplectique, on a ${}^tMJM = J$ donc $M = -J{}^tM^{-1}J$ et $\det M = 1$, d'où

$$P(\lambda) = \det(-J{}^tM^{-1}J - \lambda I) = \det({}^tM^{-1} - \lambda I) = \det(I - \lambda {}^tM) = \det(I - \lambda^t M) = \lambda^{2g}P(1/\lambda).$$

En conséquence, l'ensemble des valeurs propres de M est stables par inverse, mais $M \in \Gamma_g(n)$ donc chacune d'elles est de la forme $\lambda = 1 + n\alpha$ avec α un entier algébrique. Le groupe engendré par les valeurs propres (qui est en fait le monoïde engendré par celles-ci vu le raisonnement précédent) est donc constitué uniquement d'éléments de $\mathcal{O}_K$ congrus à 1 modulo $n\mathcal{O}_K$. Si une racine de l'unité ζ appartient à ce monoïde, on a donc $\zeta = 1 + n\alpha$ avec $\alpha \in \mathcal{O}_K$, et donc $\zeta = 1 + n\alpha$ avec α un entier algébrique de $\mathbb{Q}(\zeta)$, c'est-à-dire que n divise $1 - \zeta$ dans $\mathbb{Z}[\zeta]$, ce qui est impossible à moins que ζ soit égal à 1 et n quelconque, ou $\zeta = -1$ (et alors $n = 1$ ou 2). Ainsi, pour $n \geq 3$, le groupe $\Gamma_g(n)$ est bien net.

(*e*) Il n'y a rien à prouver, hormis que π est l'identité restreint à l'ouvert dense naturel $A_g(n)_{\mathbb{C}}$ de $A_g(n)_{\mathbb{C}}^S$ et $A_g^*(n)_{\mathbb{C}}$, donc le fibré ici nommé L prolonge bien L sur $A_g(n)_{\mathbb{C}}$. □

Dans le cas où $g = 2$, qui est notre but principal, nous allons donner une description un peu plus précise de la compactification toroïdale choisie, basée à peu de choses près sur les notations de [HW00].

Définition III.5.11 (Systèmes d'indices pour les composantes de bord)**.**

Soit $n \geq 3$ un entier. On considère le $\mathbb{Z}/n\mathbb{Z}$-module $(\mathbb{Z}/n\mathbb{Z})^4$ muni de la forme symplectique standard, c'est-à-dire celle associée à J. Ici, la notation $p|n$ signifie que p est un nombre premier divisant n.

(*a*) On note $\mathfrak{P}_1(n)$ l'ensemble des vecteurs primitifs de $(\mathbb{Z}/n\mathbb{Z})^4$ à ± 1 près, qui est de cardinal $n^4/2\prod_{p|n}(1-p^{-4})$.

(*b*) On note $\mathfrak{P}_2(n)$ l'ensemble des produits extérieurs simples $h \in \Lambda^2(\mathbb{Z}/n\mathbb{Z})^4$ à ± 1 près, où h de la forme $h = a \wedge b$ avec $(a|b)$ complétable en une matrice symplectique de $M_4(\mathbb{Z}/n\mathbb{Z})$. Cet ensemble est de cardinal $n^4/2\prod_{p|n}(1-p^{-4})$.

(*c*) On note $\mathfrak{P}_{1,2}(n)$ l'ensemble des paires (ℓ, h) avec $\ell \in \mathfrak{P}_1(n)$ et $h \in \mathfrak{P}_2(n)$ tels que ℓ appartient au $\mathbb{Z}/n\mathbb{Z}$-plan engendré par h. Cet ensemble est de cardinal $n^6/4\prod_{p|n}(1-p^{-2})(1-p^{-4})$.

Remarque III.5.4. Nous n'avons pas rédigé le décompte ci-dessus (qu'on peut retrouver dans [HW00]), mais il s'obtient par des techniques de comptage modulo n habituelles (relèvement en

puissances d'un nombre premier et théorème chinois). Il suffit donc de montrer les formules ci-dessus (sans le signe près) pour $n = p$ premier avant de diviser par deux (ou 4 pour le (c)) pour $n \geq 3$ pour tenir compte des ambiguïtés de signe.

Proposition III.5.12 (Description de $A_2^*(n)_{\mathbb{C}}$)**.**

Soit $n \geq 3$ un entier. Avec les notations précédentes et $\pi : A_2^*(n)_{\mathbb{C}} \to A_2(n)_{\mathbb{C}}^S$ le morphisme canonique :

(a) Les composantes de bord de degré 1 de $A_2(n)_{\mathbb{C}}^S$ sont paramétrées par $\mathfrak{P}_1(n)$, et on note $\mathcal{C}(\ell)$ une telle composante de bord lorsque $\ell \in \mathfrak{P}_1(n)$. De même, les composantes de bord de degré 2 de $A_2(n)_{\mathbb{C}}^S$ (qui sont des points) sont paramétrées par $\mathfrak{P}_2(n)$, et on note $\mathcal{C}(h)$ une telle composante de bord lorsque $h \in \mathfrak{P}_2(n)$.

(b) Pour tout $\ell \in \mathfrak{P}_1(n)$, l'image réciproque de $\mathcal{C}(\ell)$ par π est munie d'une identification canonique avec $\mathcal{X}_1(n)$ (voir la définition-proposition III.5.4 (b)) et d'une fibration naturelle sur $A_1(n)_{\mathbb{C}}$ induite par π. Cet isomorphisme se prolonge en un isomorphisme entre l'adhérence de $\pi^{-1}(\mathcal{C}(\ell))$ dans $A_2^*(n)_{\mathbb{C}}$, notée $D(\ell)$ (qui est un diviseur irréductible de la variété) et la surface modulaire de Shioda $S(n)$ de niveau n (c'est-à-dire la compactification canonique de la courbe elliptique universelle de niveau n, voir [Shi72], section 4). Celle-ci est munie d'un morphisme canonique vers la compactification usuelle $X(n)$ de $Y(n)_{\mathbb{C}}$ telle qu'en chaque pointe de $X(n)$, la fibre est un n-gone de droites projectives.

(c) Pour tout $h \in \mathfrak{P}_2(n)$, la fibre $\mathcal{C}^*(h) := \pi^{-1}(\mathcal{C}(h)) \subset A_2^*(n)_{\mathbb{C}}$ est une union de droites projectives disposées sous forme polyédrale qu'on peut décrire de la manière suivante. Les faces du polyèdre sont paramétrées par les $\ell \in \mathfrak{P}_1(n)$ tels que (ℓ, h) appartient à $\mathfrak{P}_{1,2}(n)$, ses arêtes (les droites projectives donc) par les paires $\{\ell_1, \ell_2\}$ d'éléments de $\mathfrak{P}_1(n)$ telles que $\ell_1 \wedge \ell_2 = \pm h$ et ses sommets par les triplets non ordonnés $\{\ell_1, \ell_2, \ell_3\}$ d'éléments de $\mathfrak{P}_1(n)$ tels que $\ell_1 \wedge \ell_2 = \pm\ell_1 \wedge \ell_3 = \pm\ell_2 \wedge \ell_3 = h$. Il y a donc respectivement $n^2/2\prod_{p|n}(1-p^{-2})$ faces, $n^3/4\prod_{p|n}(1-p^{-2})$ arêtes et $n^3/6\prod_{p|n}(1-p^{-2})$ sommets.

(d) Pour $\ell \in \mathfrak{P}_1(n)$ et $h \in \mathfrak{P}_2(n)$, si $(\ell, h) \in \mathfrak{P}_{1,2}(n)$, l'intersection de $D(\ell)$ avec $\mathcal{C}^*(h)$ est constituée des arêtes se situent sur la face paramétrée par ℓ comme ci-dessus : c'est donc un n-gone de droites projectives. Sinon, cette intersection est vide.

(e) Pour $\ell_1, \ell_2 \in \mathfrak{P}_{1,2}(n)$ distincts, l'intersection de $D(\ell_1)$ et $D(\ell_2)$ dans $A_2^*(n)_{\mathbb{C}}$ est égale à l'arête $\{\ell_1, \ell_2\}$ de $\mathcal{C}^*(h)$ si $h = \ell_1 \wedge \ell_2 \in \mathfrak{P}_2(n)$ (qu'on note alors $D(\ell_1, \ell_2)$, c'est une copie de $\mathbb{P}^1$) et vide sinon.

Démonstration. (a) C'est le cas $g = 2$ de la proposition III.5.9 (b), en prenant garde au fait que pour les composantes de degré 2, l'application qui aux deux premières colonnes d'une matrice symplectique sur $(\mathbb{Z}/n\mathbb{Z})^4$ associe leur produit extérieur induit une bijection entre l'ensemble de ces matrices à $\mathrm{GL}_2(\mathbb{Z})$ près et $\mathfrak{P}_2(n)$.

(b) L'identification de la composante de bord avec la courbe elliptique universelle de niveau n vient de la définition-proposition III.5.10 (c). Pour le prolongement à un isomorphisme entre son adhérence et la surface modulaire de Shioda, on peut le lire entre les lignes dans [Igu67a] ou bien avoir une description explicite de la compactification toroïdale ci-dessus dans [Hul00] ou la section 3 de [HW00].

(c) C'est exactement le théorème 3 de [Igu67a] dans le cas $g = 2$.

(d),(e) C'est une conséquence des paramétrisations précédentes et du (c) (voir les références déja indiquées, avec des notations légèrement différentes dans chacune).

□

Pour les besoins de la méthode de Runge, nous devons avoir de bons modèles entiers des compactifications, et c'est dans ce but que nous allons maintenant rappeler quelques résultats de [FC90].

Définition III.5.13 (Schémas abéliens et structures symplectiques de niveau)**.**

(a) Un *S-schéma abélien* est un schéma en groupes $A \to S$ lisse, propre et dont les fibres sont géométriquement connexes (c'est alors automatiquement un schéma en groupes commutatif). Il est *principalement polarisé* s'il est muni d'un isomorphisme $\lambda : A \to \widehat{A}$ (où $\widehat{A}$ est le schéma en

groupes dual) tel qu'en chaque point géométrique $\overline{s}$ de S, l'isomorphisme induit $\lambda_{\overline{s}} : A_{\overline{s}} \to \widehat{A}_{\overline{s}}$ provient d'une polarisation principale sur $A_{\overline{s}}$.

(b) Une structure symplectique de niveau $n \geq 1$ sur A principalement polarisé et de dimension relative g sur un $\mathbb{Z}[\zeta_n, 1/n]$-schéma est la donnée d'un isomorphisme symplectique $A[n] \cong (\mathbb{Z}/n\mathbb{Z})^{2g}$ pour la polarisation donnée sur A.

Définition-Proposition III.5.14. Pour tout entier $n \geq 1$:

(a) Pour tout $g \geq 1$ entier, la compactification de Satake $A_g(n)^S_{\mathbb{C}}$ a un modèle entier $\mathcal{A}_g(n)^S$ sur $\mathbb{Z}[\zeta_n, 1/n]$ qui contient comme sous-schéma ouvert dense le schéma de modules (grossier si $n < 3$) $\mathcal{A}_g(n)$ sur $\mathbb{Z}[\zeta_n, 1/n]$ paramétrant les schémas abéliens principalement polarisés munis d'une structure symplectique de niveau n. Ce modèle entier est propre et de type fini sur $\mathbb{Z}[\zeta_n, 1/n]$. Ce modèle admet également une stratification par composantes de bord qui sont indexées de la même manière que ci-dessus (y compris pour la description plus fine dans le cas $g = 2$).

(b) Pour $g \in \{2, 3\}$, la compactification toroïdale $A^*_g(n)_{\mathbb{C}}$ a un modèle entier $\mathcal{A}^*_g(n)$ sur $\mathbb{Z}[\zeta_n, 1/n]$ qui contient comme sous-schéma ouvert $\mathcal{A}_g(n)$. Ce modèle entier est propre et de type fini sur $\mathbb{Z}[\zeta_n, 1/n]$, et muni d'un morphisme surjectif canonique vers $\mathcal{A}_g(n)^S$ étendant l'identité sur $\mathcal{A}_g(n)$.

(c) Pour tout diviseur m de n, on dispose de morphismes canoniques $\mathcal{A}_g(n)^S \to \mathcal{A}_g(m)^S$ et $\mathcal{A}^*_g(n) \to \mathcal{A}^*_g(m)$ qui prolongent le morphisme de dégénerescence naturel $\mathcal{A}_g(n) \to \mathcal{A}_g(m)$.

Démonstration. Ces résultats sont contenus dans les théorèmes V.2.5 et IV.6.7 de [FC90] (attention aux notations différentes de compactification). □

III.5.2 Géométrie des compactifications

Nous avons besoin pour la méthode de Runge d'en connaître plus sur la grosseur et l'amplitude des diviseurs en jeu.

Définition III.5.15. |Groupe de Picard rationnel]

Pour toute variété normale X sur $\mathbb{C}$, le *groupe de Picard rationnel de* X est le $\mathbb{Q}$-espace vectoriel défini par $\mathrm{Pic}(X)_{\mathbb{Q}} := \mathrm{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{Q}$.

Proposition III.5.16 (Groupes de Picard de certains variétés modulaires de Siegel)**.**

Soit $g \in \{2, 3\}$ et $n \geq 1$. Alors :

(a) *Tout diviseur de Weil sur $A^*_g(n)_{\mathbb{C}}$ ou $A_g(n)^S_{\mathbb{C}}$ ou $A_g(n)_{\mathbb{C}}$ est à multiple près un diviseur de Cartier, et le groupe de Picard rationnel de ces variétés est donc le même que le groupe des diviseurs de Weil à équivalence linéaire près tensorisé par $\mathbb{Q}$.*

(b) *Les groupes de Picard rationnels de $A_g(n)^S_{\mathbb{C}}$ et $A_g(n)_{\mathbb{C}}$ sont égaux à $\mathbb{Q}L$ si $g = 3$. Pour $n = 1$, on a $\mathrm{Pic}(A^*_2(1)_{\mathbb{C}})_{\mathbb{Q}} = \mathbb{Q}L \oplus \mathbb{Q}D$.*

(c) *Pour $n = 1$, on a $\mathrm{Pic}(A_2(1)_{\mathbb{C}})_{\mathbb{Q}} = \mathrm{Pic}(A_2(1)^S_{\mathbb{C}})_{\mathbb{Q}} = \mathbb{Q}L$ et $\mathrm{Pic}(A_2(1)^*_{\mathbb{C}}) = \mathbb{Q}L \oplus \mathbb{Q}D$.*

Démonstration. (a) Tout d'abord, ce résultat est vrai pour $A^*_g(n)_{\mathbb{C}}$ et $A_g(n)_{\mathbb{C}}$ car celles-ci sont à singularités de quotient finies d'après [ABMMOG14] (le résultat principal de cet article semble avoir été admis dans la littérature bien avant...). Maintenant, comme le complémentaire de $A_g(n)_{\mathbb{C}}$ dans $A_g(n)^S_{\mathbb{C}}$ est de codimension au moins 2, les groupes des diviseurs de Weil à équivalence linéaire près de ces deux variétés sont égaux, or le premier est à $\mathbb{Q}$ près le groupe de Picard, donc le second également, ce qui prouve le (a).

(b) Pour $g = 3$ et la compactification de Satake, c'est une conséquence de résultats généraux de Borel [Bor81] que le groupe de Picard rationnel pour $A_g(n)_{\mathbb{C}}$ ou sa compactification ne change pas lorsque n augmente [Wei92], et reste égal à $\mathbb{Q}L$ (c'est même vrai pour tout $g \geq 3$). On en déduit le résultat sur $A^*_3(1)$ car D est irréductible lorsque $n = 1$, et par ([Har77], Proposition II.6.5).

Pour le cas $g = 2$ et la compactification toroïdale : c'est une conséquence du théorème 2 de [vdG98] (le résultat sur le groupe de Picard en lui-même est plus ancien que ceci, voir la partie III de [Mum83]). On en déduit ensuite le groupe de Picard sur la partie ouverte par la proposition II.6.5 de [Har77] (encore une fois car le bord D est irréductible et de codimension 1), puis sur la compactification de Satake. □

Remarque III.5.5. Il semble beaucoup plus difficile de calculer le groupe de Picard rationnel de $A_2(n)_{\mathbb{C}}$ pour n grand : des approches en sont données par exemple dans [Wei92] (générateurs du groupe pour tout n, mais pas la dimension sur $\mathbb{Q}$ ni les relations) ou [HW00] (pour $n = 3$, avec cette fois une description explicite d'une $\mathbb{Q}$-base).

En fonction de ces générateurs, nous pouvons maintenant caractériser les diviseurs amples sur ces diverses variétés de la manière suivante.

Proposition III.5.17 (Diviseurs amples sur les variétés modulaires de Siegel)**.**
Pour $g \in \{2, 3\}$:
(a) Un $\mathbb{Q}$-diviseur aL est ample dans $A_3(n)_{\mathbb{C}}$, dans $A_3(n)^S_{\mathbb{C}}$, dans $A_2(1)_{\mathbb{C}}$ ou dans $A_2(1)^S_{\mathbb{C}}$ si et seulement si $a > 0$.
(b) Un diviseur $aL - bD$ est ample dans $A^*_g(n)_{\mathbb{C}}$ si et seulement si $b > 0$ et $a > 12b/n$.

Démonstration. Rappelons que dans une variété projective X, le $\mathbb{R}$-espace vectoriel $N^1(X)$ défini par $N^1(X) = \mathrm{NS}(X) \otimes_{\mathbb{Z}} \mathbb{R}$ (où $\mathrm{NS}(X)$ est le groupe de Néron-Severi de X) est de dimension finie (c'est le théorème de la base, proposition 1.1.14 de [Laz04]), et l'ensemble des $\mathbb{R}$-diviseurs nef de X (c'est-à-dire des sommes formelles à coefficients réels positifs de diviseurs dont le nombre d'intersection avec toute courbe irréductible de X est positif, voir la section 1.4.2 de [Laz04]) forme un cône de $N^1(X)$ dont l'intérieur est exactement l'ensemble des $\mathbb{R}$-diviseurs amples de X ([Laz04], théorème 1.4.21). Il suffit donc de connaître les résultats du (a) et du (b) avec des inégalités larges pour le cône nef pour en déduire le cône ample donc le (a) et le (b).
Pour le (a), c'est dû à la construction de L, voir les propositions III.5.16 et III.5.8 (d).
Pour le (b), c'est le théorème 0.1 et le théorème 0.2 de [Hul00]. □

Remarque III.5.6. Le fait que l'inégalité $a > 12b/n$ semble être moins exigeante n'est pas une contradiction, car le morphisme d'oubli $A^*_g(n)_{\mathbb{C}} \to A^*_g(1)_{\mathbb{C}}$ est ramifié de degré n le long du diviseur du bord, d'où cette équivalence des conditions.

Comme la méthode de Runge accepte également des gros diviseurs, rappelons quels sont les gros diviseurs sur ces variétés.

Proposition III.5.18 (Diviseurs gros sur les variétés modulaires de Siegel)**.**
(a) Un $\mathbb{Q}$-diviseur aL est gros dans $A_3(n)_{\mathbb{C}}$, dans $A_3(n)^S_{\mathbb{C}}$ ou dans $A_2(1)_{\mathbb{C}}$ ou $A_2(1)^S_{\mathbb{C}}$ si et seulement si $a > 0$.
(b) Un $\mathbb{Q}$-diviseur $aL - bD$ est gros dans $A^*_2(1)_{\mathbb{C}}$ si et seulement si $a > 0$ et $b < a/10$. Si de plus $b < a/12$, le $\mathbb{Q}$-diviseur $aL - bD$ induit un morphisme $A^*_2(1)_{\mathbb{C}} \to \mathbb{P}^k$ qui est un plongement hors du support de D.

Démonstration. Sur une variété normale projective X, un diviseur est gros si et seulement à un multiple entier naturel près, c'est la somme d'un diviseur ample A et d'un diviseur effectif E ([Laz04], Corollaire 2.2.6). Dans le cas du (a), les diviseurs effectifs sur les différentes variétés considérées sont tous de la forme $aL, a \geq 0$ (proposition III.5.16), donc les diviseurs amples sont les mêmes que les diviseurs gros.

Pour le (b), la situation est plus complexe, et il faut donc d'abord savoir quels sont les $aL - bD$ qui sont linéairement équivalents à des $\mathbb{Q}$-diviseurs effectifs dans $A^*_2(1)_{\mathbb{C}}$. L'application de Torelli $t : \overline{\mathcal{M}_2} \to A^*_2(1)_{\mathbb{C}}$, où $\overline{\mathcal{M}_2}$ est l'espace des courbes stables de genre 2, est surjective finie, et donc pour tout diviseur effectif sur $A^*_2(1)_{\mathbb{C}}$, son pullback par t est aussi un diviseur effectif. Le problème des diviseurs effectifs dans $A^*_2(1)_{\mathbb{C}}$ devient donc ce qu'on appelle un « problème de pente » dans $\overline{\mathcal{M}_2}$, à savoir trouver pour un diviseur $a\lambda - b\delta_1$, où λ est le fibré en droites de Hodge et δ_1 le diviseur du bord de $\overline{\mathcal{M}_2}$ envoyé par t sur D, le rapport minimal a/b si $b > 0$. Ce problème de pente est un champ d'études en soi (généralisé à g quelconque pour les variétés modulaires de Siegel ou les espaces de modules de courbes, voir [CFM13]), mais on sait que le rapport minimal a/b vaut 10 pour notre problème ici (théorème 0.4 de [HM90], nous avons repris les notations de cet article dans ce paragraphe). En conséquence, comme $t^*L = \lambda$ et $t^*D = \delta_1$, si un diviseur linéairement équivalent à $aL - bD$ est effectif avec $b > 0$, alors $a \geq 10b$ (nous verrons plus tard

quel diviseur réalise l'égalité). On voit également aisément que si $aL - bD$ est effectif, $a \geq 0$ pour des raisons d'intersection, et donc le cône pseudoeffectif de $A_2^*(1)_{\mathbb{C}}$ (c'est-à-dire engendré par les diviseurs effectifs) est exactement constitué des diviseurs $aL - bD$ tels que $a \geq 0$ et $a \geq 10b$ (en particulier, pour a nul, on retrouve exactement les multiples positifs de D).

D'après ([Laz04], Théorème 2.2.24), un diviseur est gros si et seulement s'il appartient à l'intérieur du cône pseudoeffectif , et dans notre cas on en déduit donc que $aL - bD$ est gros si et seulement si $a > 0$ et $a > 10b$. Remarquons de plus que si $a > 0$ et $b < a/12$, si b est positif le diviseur est ample, et sinon $aL - bD = aL - (a/13)D + b'D$ avec $b' > 0$ donc $aL - bD = A + b'D$ avec A ample et $b'D$ un $\mathbb{Q}$-diviseur effectif.

Alors, (quitte à avoir multiplié pour que A soit très ample et $b'D$ un vrai diviseur effectif), si $s_0, \cdots, s_n$ est une base de sections du fibré associé à A sur $A_2^*(1)_{\mathbb{C}}$ et s_D une section non nulle du fibré associé à $b'D$ (correspondant par exemple à la fonction constante 1 sur la variété comme $b'D$ est effectif), les $s_0 \otimes s_D, \cdots, s_n \otimes s_D$ sont des sections du fibré associé au diviseur $A + b'D$ qui définissent une application rationnelle hors du support de D, qui coïncide avec le plongement associé aux sections $s_0, \cdots, s_n$ hors de ce support. Ainsi, le morphisme induit par une base de sections d'un multiple assez grand du fibré associé à $A + b'D$ est un plongement hors du support de D, ce qu'on voulait démontrer. □

Nous savons maintenant quels diviseurs sont gros et amples sur les variétés modulaires de Siegel qu'on souhaite étudier, il est temps de fournir des exemples précis de diviseurs candidats à la méthode de Runge.

III.5.3 Un théorème à la Runge pour la torsion des diviseurs thêta de surfaces abéliennes

Définition III.5.19 (Diviseur thêta sur une surface abélienne)**.**

Soit k un corps algébriquement clos. Soit A une surface abélienne définie sur k, et L un fibré en droites ample symétrique sur A induisant une polarisation principale $\lambda : A \to \widehat{A}$ (où $\widehat{A}$ dénotera toujours par la suite la variété abélienne duale de A).

Alors, $h^0(A, L)$ est un k-espace vectoriel de dimension 1 ([Mum86], Théorème de Riemann-Roch p. 150 appliqué dans le cas ample). Une *fonction thêta associée à* (A, L) est une section globale non nulle $\vartheta_{A,L}$ de L, et le *diviseur thêta associé à* (A, L), noté $\Theta_{A,L}$, est le diviseur des zéros de $\vartheta_{A,L}$ (indépendant de notre choix).

Le diviseur thêta sur A est déterminé par la polarisation principale associée à L, à une ambiguïté finie près qu'on va expliciter ci-dessous.

Proposition III.5.20. *Soit k un corps algébriquement clos et A une variété abélienne de genre g définie sur k.*

Deux fibrés amples symétriques L et L' induisent la même polarisation principale si et seulement si $L' = T_x^* L$ pour un certain point $x \in A(k)$ de 2-torsion, et alors $\Theta_{A,L'} = \Theta_{A,L} + x$.

Démonstration. Pour tout fibré L sur A, on note $\lambda_L : A \to \widehat{A} \cong \mathrm{Pic}^0(A)$ l'application qui à x associe $T_x^* L \otimes L^{-1}$. C'est un morphisme de groupes et l'application qui à L associe λ_L est additive de $\mathrm{Pic}(A)$ dans $\mathrm{Hom}(A, \widehat{A})$, et de noyau $\mathrm{Pic}^0(A)$ ([Mum86], Chapitre II, Corollaire 4 et ce qui suit, et section II.8). De plus, lorsque L est ample, cette application est une isogénie ([Mum86], Chapitre II, Théorème 1, c'est en fait la polarisation associée à L), et donc si L et L' sont amples. En particulier, pour tout $x \in A(k)$, si $L' \cong T_x^* L$, alors $L' \otimes L^{-1} \in \mathrm{Pic}^0(A)$, donc L' induit la même polarisation principale que L si celui-ci en induit une. Réciproquement, si $\lambda_L = \lambda_{L'}$ pour deux fibrés amples, alors $L' \otimes L^{-1}$ appartient à $\mathrm{Pic}^0(X)$, et donc par surjectivité il existe $x \in A(k)$ tel que $L' \otimes L^{-1} \cong T_x^* L \otimes L^{-1}$ d'où $L' \cong T_x^* L$. De plus, si L et L' sont symétriques, on a $[-1]^* L \cong L$ et donc $[-1]^* L' \cong T_{-x}^* [-1]^* L \cong T_{-x}^* L$ donc $T_{-x}^* L \cong T_x^* L$ mais si la polarisation induite par L est principale, le morphisme λ_L est un isomorphisme, et alors $-x = x$ c'est-à-dire que $x \in A[2]$.

Alors, pour $\vartheta_{A,L}$ une section non nulle de L, son pullback $T_x^* \vartheta_{A,L}$ est une section non nulle d'un fibré isomorphe à L', et donc $\Theta_{A,L'} = \Theta_{A,L} - x = \Theta_{A,L} + x$ car $-x = x$. □

En caractéristique différente de 2, en ajoutant à (A, λ) la donnée d'une 2-structure symplectique, on peut établir une détermination complète de L ample symétrique associé à λ selon le procédé suivant qu'on appellera la *correspondance d'Igusa* ([Igu67b], Théorème 2 et ce qui le précède). À tout diviseur de Weil ample symétrique D sur A définissant une polarisation principale, on fait correspondre bijectivement une forme quadratique q_D sur $A[2]$ à valeurs dans $\{\pm 1\}$ dite paire (c'est-à-dire que la somme de ses valeurs sur $A[2]$ vaut 2^g). Or, une fois fixée une base symplectique $\mathcal{B}_2$ de $A[2]$, l'application q qui à $x \in A[2]$ de coordonnées $(m', m'') \in (\mathbb{Z}/2\mathbb{Z})^{2g}$ dans $\mathcal{B}_2$ associe

$$q(x) = (-1)^{m'^t m''}$$

est bien une forme quadratique paire sur $A[2]$, et on choisit alors comme diviseur ample symétrique l'unique D correspondant à q. De plus, par construction de cette correspondance (voir [Igu67b], p. 823), un point x de 2-torsion de A de coordonnées (a, b) dans $\mathcal{B}_2$ appartient au diviseur thêta correspondant à l'unique fibré ample symétrique associé à $\mathcal{B}_2$ si $a^t b = 1$ modulo 2. Pour un point de coordonnées (a, b) dans $\mathcal{B}_2$ avec $a^t b = 0$ modulo 2, s'il appartient au diviseur thêta, c'est avec multiplicité paire.

Définition-Proposition III.5.21 (Diviseur Thêta canoniquement associé à une n-structure symplectique)**.**

Soit $n \geq 2$ pair et k un corps algébriquement clos de caractéristique ne divisant pas n.

Pour un triplet (A, λ, α_n) avec A une variété abélienne de dimension g, λ une polarisation principale sur A et α_n une structure symplectique de niveau n sur A (définition III.5.13), il existe un unique fibré ample symétrique L sur A induisant λ et correspondant à la base symplectique de $A[2]$ induite par α_n comme ci-dessus. On appelle alors le diviseur Θ_{A,L,α_n} correspondant « le diviseur thêta » associé à ce triplet.

Nous cherchons dans cette section à donner un résultat à la Runge pour les surfaces abéliennes principalement polarisées (A, λ) sur un corps de nombres K dont le diviseur thêta ne contient aucun point de n-torsion de A sauf des points de 2-torsion. On verra dans la proposition III.5.25 que cela implique que (A, λ) n'est pas un produit de courbes elliptiques, mais cette condition nécessaire n'est pas suffisante (cf. par exemple [BG00]).

Ces notions sont intimement liées aux fonctions thêta complexes classiques, comme nous allons le démontrer ci-dessous.

Définition-Proposition III.5.22 (Fonctions thêta complexes)**.**

Dans toutes les notations ci-dessous, les éléments de $\mathbb{Z}^g, \mathbb{R}^g$ ou $\mathbb{C}^g$ sont des vecteurs lignes.

On définit la fonction holomorphe Θ sur $\mathbb{C}^g \times \mathcal{H}_g$ par la série convergente sur tout compact

$$\Theta(z, \tau) = \sum_{n \in \mathbb{Z}^g} e^{i\pi n \tau^t n + 2i\pi n^t z}.$$

Pour tous $a, b \in \mathbb{R}^g$, on définit également les fonctions $\Theta_{a,b}$ par

$$\Theta_{a,b}(z, \tau) = \sum_{n \in \mathbb{Z}^g} e^{i\pi (n+a)\tau^t (n+a) + 2i\pi (n+a)^t (z+b)}.$$

Pour $\tau \in \mathcal{H}_g$ fixé, on note Θ_τ la fonction $z \mapsto \Theta(z, \tau)$ et de même pour $\Theta_{a,b,\tau}$.

(a) Pour tous $p, q \in \mathbb{Z}^g$, la fonction Θ vérifie l'équation fonctionnelle

$$\Theta(z + p\tau + q) = e^{-i\pi p \tau^t p - 2i\pi p^t z} \Theta(z, \tau).$$

Les fonctions $\Theta_{a,b}$ vérifient

$$\Theta_{a,b}(z, \tau) = e^{i\pi a \tau^t a + 2i\pi a^t (z+b)} \Theta(z + a\tau + b, \tau),$$

et de plus

$$\Theta_{a,b}(z + p\tau + q) = e^{-i\pi p \tau^t p - 2i\pi p^t z + 2i\pi (a^t q - b^t p)} \Theta_{a,b}(z, \tau).$$

(*b*) On désigne par ϑ et $\vartheta_{a,b}$ les « thêta-constantes normalisées », à savoir les fonctions holomorphes sur $\mathcal{H}_g$ définies par

$$\vartheta(\tau) := \Theta(0,\tau) \quad \text{et} \quad \vartheta_{a,b}(\tau) := e^{-i\pi a^t b}\Theta_{a,b}(0,\tau).$$

Alors, ces fonctions thêta vérifient la propriété de modularité suivante : pour tout $\gamma \in \Gamma_g(2)$,

$$\vartheta_{a,b}(\gamma\cdot\tau) = \zeta_8(\gamma)e^{i\pi(ab)^t V_\gamma}\sqrt{j_\gamma(\tau)}\vartheta_{(a,b)M}(\tau), \tag{III.7}$$

où $\zeta_8(\gamma)$ (qui est une racine huitième de l'unité) et $V_\gamma \in \mathbb{Z}^g$ ne dépendent que de γ, et avec les notations de la définition III.5.1.

En particulier, pour un entier $n \geq 2$ pair, si $na, nb \in \mathbb{Z}^g$, la fonction $\vartheta_{a,b}^{8n}$ est une forme modulaire de Siegel de degré g, de niveau n, et de poids $4n$ qui dépend seulement de $(a,b) \mod \mathbb{Z}^{2g}$.

Démonstration. La convergence de ces séries et leurs équations fonctionnelles sont l'objet de ([Mum87], section II.1) (où les vecteurs sont notés en colonnes). Pour la propriété de modularité, c'est un cas particulier des calculs de la section II.5 de [Mum87] (on n'aura pas besoin de la formule générale). Ensuite, on observe par manipulation des séries définissant Θ que pour tous $p, q \in \mathbb{Z}^g$, on a

$$\vartheta_{a+p,b+q} = e^{2i\pi(a^t q - b^t p}\vartheta_{a,b}.$$

En conséquence, si $(na, nb) \in \mathbb{Z}^{2g}$, $\vartheta_{a,b}^n$ dépend seulement de $(a,b) \mod \mathbb{Z}^{2g}$. En passant la propriété de modularité à la puissance $8n$, on neutralise également la racine de l'unité provenant de la propriété de modularité, donc $\vartheta_{a,b}^{8n}$ est bien une forme modulaire de Siegel de poids $4n$ pour $\Gamma_g(n)$. □

Remarque III.5.7. Dans le cas $g = 1$, nous avons prouvé dans une section précédente, et de manière totalement explicite, la formule (III.7) avec des normalisations légèrement différentes (théorème III.5).

Les notions de diviseur thêta et les fonctions thêta ci-dessus sont liées de la manière suivante.

Proposition III.5.23.
On reprend les notations de la définition-proposition III.5.3.
Soit $\tau \in \mathcal{H}_g$. Soit L_τ le fibré en droites complexes sur $A_\tau = \mathbb{C}^g/\Lambda_\tau$ défini comme le quotient de $\mathbb{C}^g \times \mathbb{C}$ par l'action de Λ_τ définie par

$$(p\tau + q)\cdot(z,t) := (z + p\tau + q, e^{-i\pi p\tau^t p - 2i\pi p^t z}t).$$

Ce fibré en droites est ample et symétrique sur A_τ, et la polarisation complexe associée à L_τ est exactement ω_τ donc principale. Réciproquement, selon la correspondance d'Igusa (définition-proposition III.5.21), le fibré en droites ample symétrique associé à ω_τ et la 2-structure canonique $\alpha_{\tau,2}$ sur A_τ est exactement L_τ.

De plus, les sections globales de L_τ s'identifient canoniquement aux multiples de Θ_τ, de sorte que le diviseur thêta associé à (A_τ, ω_τ) est exactement le diviseur des zéros de Θ_τ modulo Λ_τ.

Ainsi, pour tous $a, b \in \mathbb{R}^g$, l'image de $a\tau + b$ dans A_τ appartient au support de $\Theta_{A_\tau,\omega_\tau,\alpha_{\tau,2}}$ si et seulement si $\vartheta_{a,b}(\tau) = 0$.

Démonstration. La vérification de la bonne définition de l'action et du fibré en droites est immédiate. Pour la polarisation principale, on retrouve comme forme bilinéaire associée au facteur de multiplication la forme de Riemann ω_τ, qui correspond bien à une polarisation principale. On observe que L_τ est symétrique, en remarquant que le pullback $[-1]^*L_\tau$ peut en fait également se voir comme un quotient de $\mathbb{C}^g \times \mathbb{C}$ par une action de Λ_τ, qui se trouve être exactement la même. Ensuite, les sections globales de L_τ se relèvent par le morphisme quotient $\mathbb{C}^g \times \mathbb{C} \to L_\tau$ en des fonctions $z \mapsto (z, f(z))$ et les fonctions f ainsi obtenues sont exactement celles qui vérifient la même propriété de multiplicativité (pour translation par Λ_τ) que la fonction Θ_τ, ce sont donc ses

multiples. Cette identification est également compatible avec les diviseurs associés, donc Θ_{A_τ,L_τ} est bien le diviseur des zéros de Θ_τ modulo Λ_τ. Pour plus de détails ou de généralités sur l'une des affirmations utilisées ici (et les fibrés en droites complexes sur A_τ), voir ([Deb99], Chapitres IV, V et section VI.2).

Il reste maintenant à vérifier que la correspondance d'Igusa associe bien L_τ à A_τ muni de sa 2-structure canonique. En regardant de près la construction de cette correspondance ([Igu67b], pp. 822,823 et 833), la fonction méromorphe ψ_x sur A_τ (dépendant de L_τ) associée à $x = a\tau+b \in A_\tau[2]$ a pour diviseur $[2]^*T_x^*\Theta_{A_\tau,L_\tau} - [2]^*\Theta_{A_\tau,L_\tau}$, c'est donc exactement le passage au quotient par Λ_τ de la fonction méromorphe f_x sur $\mathbb{C}^g$ définie par

$$f_x(z) = \frac{\Theta_{a,b,\tau}(2z)}{\Theta_\tau(2z)}$$

(cette fonction est bien invariante par translation par Λ_τ grâce à la définition-proposition III.5.22 (a)). Ensuite, la forme quadratique q de $A[2]$ dans $\{\pm 1\}$ correspondante par Igusa est définie par l'identité

$$f_x(-z) = q(x)f_x(z)$$

pour tout $z \in \mathbb{C}^g$, or Θ_τ est paire, donc

$$f_x(-z) = e^{2i\pi a^t b} f_x(z)$$

d'après les formules de la définition-proposition III.5.22 (a). Or, les coordonnées de la base canonique de $A[2]$ de l'image de x sont exactement $(2b \mod 2\mathbb{Z}^g, 2a \mod 2\mathbb{Z}^g)$ par définition, donc le facteur de parité est bien celui voulu pour le fibré ample symétrique associé à la 2-structure (voir la section 3 de [Igu67b] pour plus de détails).

Faisons maintenant le lien entre les zéros des thêta-constantes et les diviseurs thêta : par le raisonnement ci-dessus, le diviseur des zéros de Θ_τ modulo Λ_τ est exactement Θ_{A_τ,L_τ}, donc pour tout $z \in \mathbb{C}^g$, on a $\Theta_\tau(z) = 0$ si et seulement si z modulo Λ_τ appartient à Θ_{A_τ,L_τ}, d'où le lien entre $\vartheta_{a,b}$ et l'appartenance de $a\tau + b$ au diviseur thêta associé à $(A_\tau, \omega_\tau, \alpha_{\tau,2})$ par la formule de la définition-proposition III.5.22 (a). □

Voyons maintenant comment on peut décrire le diviseur thêta associé à (E, L) (sur un corps algébriquement clos quelconque) lorsque E est une courbe elliptique.

Proposition III.5.24. *Soit E une courbe elliptique sur le corps algébriquement clos k de caractéristique différente de 2, et L un fibré ample symétrique définissant une polarisation principale sur E. Alors, le diviseur effectif $\Theta_{E,L}$ est un point de 2-torsion de E sans multiplicité. De plus, si E est muni d'une base (e_1, e_2) de sa 2-structure et L le fibré ample symétrique associé par la correspondance d'Igusa, alors le point en question est exactement $e_1 + e_2$.*

Démonstration. Par le théorème de Riemann-Roch sur E, le diviseur effectif $\Theta_{E,L}$ est de degré 1 car $h^0(E, L) = 1$, il est donc réduit à un point x de E sans multiplicité. De plus, ce diviseur est stable par $[-1]$ car L est symétrique, donc $x = -x$ (rappelons que Riemann-Roch prouve également que deux points distincts de E ne peuvent avoir de diviseurs associés linéairement équivalents). Maintenant, d'après la construction d'Igusa, par imparité de la fonction $f_{ae_1+be_2}$ définie ci-dessus si $ab = 1$, on a $x = e_1 + e_2$. □

Remarque III.5.8. Sur $\mathbb{C}$, ce résultat couplé avec la proposition III.5.23 prouve point par point que la thêta-constante $\vartheta_{\frac{1}{2},\frac{1}{2}}$ est toujours nulle sur $\mathcal{H}_1$ mais que les autres $\vartheta_{a,b}$ avec $0 \leq a < 1$ et $0 \leq b < 1$ ne s'annulent jamais sur $\mathcal{H}_1$: on avait déja obtenu ces résultats de manière indépendante avec le lemme III.3.8. Nous allons voir dans les propositions III.5.25 et III.5.27 que la situation est loin d'être aussi simple pour les $\vartheta_{a,b}$ où $a, b \in \mathbb{Q}^g$.

Ceci nous permet de décrire le diviseur thêta d'un produit de courbes elliptiques.

Proposition III.5.25. *Soit (A, L) avec $A = E_1 \times E_2$ un produit de courbes elliptiques sur le corps algébriquement clos k de caractéristique différente de 2, et L ample symétrique sur E, induisant la polarisation principale produit sur A. Alors, le diviseur $\Theta_{A,L}$ est de la forme*

$$\Theta_{A,L} = \{x_1\} \times E_2 + E_1 \times \{x_2\}$$

avec x_i un point de 2-torsion de E_i pour $i = 1, 2$. En particulier, il admet un point singulier en (x_1, x_2), et pour tout $n \in \mathbb{N}$ pair, et non nul dans k :

(a) Le nombre de points de 2-torsion de $\Theta_{A,L}$ est exactement 7.

(b) Le nombre de points de n-torsion (mais pas de 2-torsion) de $\Theta_{A,L}$ est exactement $2(n^2-4)$.

Démonstration. Une section de (A, L) s'écrivant comme produit tensoriel de sections de E_1 et E_2 pour leur polarisation principale, la structure du diviseur $\Theta_{A,L}$ est une conséquence directe de la proposition III.5.24. Ensuite, le nombre de points de n-torsion de $\{x_1\} \times E_1$ (ou de $E_2 \times \{x_2\}$) est exactement n^2 pour tout n pair car la caractéristique de k ne divise pas n, d'où les comptages du (a) et (b) car l'intersection de ces deux ensembles est (x_1, x_2), qui est de 2-torsion. □

Pour comprendre les autres cas de figure en dimension 2, rappelons un résultat fondamental sur les surfaces abéliennes.

Proposition III.5.26. *Soit k un corps quelconque.*

Une surface abélienne principalement polarisée (A, λ) sur k est, après éventuelle extension finie des scalaires, soit un produit de courbes elliptiques soit la jacobienne J d'une courbe lisse C de genre 2 (dans les deux cas munie de sa polarisation principale canonique), et dans le second cas, pour un plongement quelconque $\phi : C \to J$ de point-base x et un fibré ample symétrique L sur J associé à la polarisation principale, le diviseur $\Theta_{J,L}$ est irréductible et c'est en fait un translaté de l'image de C par ϕ.

Démonstration. Cette proposition (et le cas de la dimension 3) sont l'objet de [OU73] (à noter que la preuve dans cet article part du cas complexe avant d'obtenir le résultat pour tous les autres corps par des techniques de descente et de schémas). □

Sauf mention contraire, on raisonne dans la suite sur un corps algébriquement clos et de caractéristique différente de 2.

Soit C une courbe hyperelliptique de genre 2 et ι son involution hyperelliptique. Alors, la courbe C a six points de Weierstrass (les points fixes de ι), et si on choisit l'un d'entre eux noté ∞, pour le morphisme d'Albanese

$$\phi : \left| \begin{array}{ccc} C & \longrightarrow & J = \mathrm{Jac}(C) \\ x & \longmapsto & [x] - [\infty] \end{array} \right. ,$$

alors $\phi(C)$ est stable par $[-1]$, car le diviseur $[x] + [\iota(x)] - 2[\infty]$ est principal pour tout $x \in C$. En conséquence, comme $\Theta_{J,L}$ est également symétrique et que c'est un translaté de $\phi(C)$, on a $\Theta_{J,L} = T_x^*(\phi(C))$ avec x un point de 2-torsion de J.

Comprendre les points du diviseur thêta revient donc à comprendre la courbe elle-même lorsqu'elle est plongée dans sa jacobienne (et en particulier la structure additive sous-jacente).

C'est un problème difficile de savoir quels points de torsion de la surface abélienne appartiennent au diviseur thêta (voir [BG00]), mais ici nous avons simplement besoin de majorer leur nombre, dans le cas des jacobiennes. La proposition ci-dessous prouve qu'il ne peut pas trop y en avoir.

Proposition III.5.27. *Soit C une courbe lisse de genre 2 définie sur le corps algébriquement clos k de caractéristique différente de 2, et (J, λ) sa jacobienne canoniquement polarisée.*

Soit ∞ un point de Weierstrass fixé de C, qui induit un plongement de C dans J par l'application $x \mapsto [x] - [\infty]$, dont on note l'image $\widetilde{C}$.

Alors, l'ensemble $\widetilde{C} \subset J$ contient 0 et est stable par multiplication par $[-1]$, et l'application

$$\begin{array}{ccc} \mathrm{Sym}^2(\widetilde{C}) & \longrightarrow & J \\ \{P, Q\} & \longmapsto & P + Q \end{array}$$

est injective en-dehors de la fibre au-dessus de 0.

En conséquence :

(*a*) *Le cardinal de* $\widetilde{C} \cap J[2](k)$ *est exactement 6.*

(*b*) *Pour tout entier pair* n *non nul dans* k*, le cardinal de* $\widetilde{C} \cap J[n](k)$ *est au plus* $\sqrt{2}n^2$*.*

Démonstration. Cette proposition n'est pas un résultat nouveau, on peut le trouver (avec des formulations légèrement différentes) dans [BG00] (Théorème 1.3) ou bien dans [Paz13] (lemme 5.1), présenté comme une conséquence du théorème d'Abel-Jacobi sur $\mathbb{C}$. Nous allons ici fournir une preuve plus détaillée.

Commençons par noter que comme $[\infty]$ est un point de Weierstrass, le diviseur $2[\infty]$ est canonique. Réciproquement, si un diviseur D de degré 2 vérifie $\ell(D) := \dim H^0(C, \mathcal{O}_C(D)) \geq 2$, alors il est canonique d'après le théorème de Riemann-Roch : en effet, on a $\ell(2[\infty] - D) \geq 1$, mais comme $2[\infty] - D$ est de degré 0, cela signifie qu'il existe une fonction f sur C de diviseur exactement $2[\infty] - D$, donc D est lui-même un diviseur canonique. Maintenant, soient quatre points x, y, z, t de C tels que $[x] + [y] = [z] + [t]$ dans J. Alors, le diviseur $[x] + [y] - [z] - [t]$ est le diviseur d'une certaine fonction f, mais cela signifie soit que f est constante (donc $\{x, y\} = \{z, t\}$) soit que $\ell([z] + [t]) \geq 2$, et d'après le raisonnement ci-dessus $[z] + [t]$ est donc un diviseur canonique. Ainsi, on a $[z] + [t] - 2[\infty] = 0$ dans J c'est-à-dire que les points $P = [x] - [\infty]$ et $Q = [y] - [\infty]$ de $\widetilde{C}$ sont de somme nulle.

En notant $\widetilde{C}[n] = \widetilde{C} \cap J[n](k) \subset J(k)$, l'application somme de $\widetilde{C}[n]^2$ dans $J[n](k)$ a une fibre de cardinal $|\widetilde{C}[n]|$ au-dessus de 0 et de cardinal au plus 2 au-dessus de tous les autres points de $J[n]$ par l'argument précédent, d'où l'inégalité du second degré

$$|\widetilde{C}(n)|^2 \leq |\widetilde{C}(n)| + 2(n^4 - 1),$$

à partir de laquelle on obtient directement l'estimation (*b*) de la proposition.

Pour le (*a*), il suffit de voir que si pour $x \in C$, $2[x]$ est linéairement équivalent à $2[\infty]$, alors $2[x]$ est canonique donc x est un point de Weierstrass par définition, et ceux-ci sont au nombre de 6. □

Nous pouvons maintenant définir les diviseurs nécessaires à notre utilisation du théorème de Runge tubulaire (théorème III.8).

Définition-Proposition III.5.28 (Diviseurs thêta sur $A_2(n)^S_{\mathbb{C}}$)**.**

Soit $n \in \mathbb{N}^*$ pair.

(*a*) On dit que $(a, b) \in (\mathbb{Z}/n\mathbb{Z})^4$ *vérifie la condition* $(*)$ si $(2a, 2b) \neq (0, 0)$ ou si $(-1)^{4a^t b/n^2} = 1$. Il y a donc exactement 6 couples $(a, b) \in (\mathbb{Z}/n\mathbb{Z})^4$ ne vérifiant pas cette condition.

(*b*) Si $(a, b) \in (\mathbb{Z}/n\mathbb{Z})^4$ vérifie la condition $(*)$, pour tout relèvement $(\tilde{a}, \tilde{b}) \in \mathbb{Z}^4$ de (a, b) dans $\mathbb{Z}^4$, la fonction $\vartheta^{8n}_{\tilde{a}/n, \tilde{b}/n}$ est une forme modulaire de Siegel non nulle de degré 2, de poids $4n$ et de niveau n, indépendante du choix des relèvements. Elle définit donc un diviseur effectif de Weil réduit sur $A_2(n)^S_{\mathbb{C}}$, appelé *diviseur thêta associé à* (a, b) et noté $(D_{n,a,b})_{\mathbb{C}}$.

(*c*) Pour (a, b) et (a', b') vérifiant la condition $(*)$, les diviseurs $(D_{n,a,b})_{\mathbb{C}}$ et $(D_{n,a',b'})_{\mathbb{C}}$ sont égaux si et seulement si $(a, b) = \pm(a', b')$. L'ensemble des paires vérifiant la condition $(*)$ définit donc exactement $n^4/2 + 2$ diviseurs thêta distincts deux à deux.

Démonstration. (*a*) Soit $(a, b) \in (\mathbb{Z}/n\mathbb{Z})^4$ tel que $(2a, 2b) = (0, 0)$. On peut alors l'écrire sous la forme $(a, b) = (na'/2, nb'/2) \mod n\mathbb{Z}^4$ avec $a', b' \in \mathbb{Z}^2$. En conséquence, $(-1)^{4a^t b/n^2} = (-1)^{a'^t b'}$, et il suffit donc de connaître les solutions modulo 2 de $(-1)^{a'^t b'} = -1$. On observe directement que celles-ci sont au nombre de six, à savoir $(0, 1, 0, 1), (1, 0, 1, 0), (1, 1, 0, 1), (1, 1, 1, 0), (1, 0, 1, 1)$ et $(0, 1, 1, 1)$.

(*b*) Grâce à la définition-proposition III.5.22 (*b*), il suffit de montrer que pour tout $(a, b) \in \mathbb{Q}^4$ tel que $(na, nb) \in \mathbb{Z}^4$, la fonction holomorphe $\vartheta_{a,b}$ n'est pas la fonction nulle à moins que $(2a, 2b)$ soit entier et ne vérifie pas la condition $(*)$ (c'est-à-dire l'un des six cas précédents). On va utiliser le développement en série de Fourier de cette forme modulaire, qu'on explique ci-dessous (pour plus de détails sur les développements de Fourier de formes modulaires de Siegel, voir la partie 4 de [Kli90]).

Pour toute matrice symétrique $S \in M_2(\mathbb{Q})$ telle que $S/(2n^2)$ est demi-entière (c'est-à-dire avec des coefficients entiers sur la diagonale, et demi-entiers ou entiers ailleurs), on a $\vartheta_{a,b}(\tau+S) = \vartheta_{a,b}(\tau)$ pour tout $\tau \in \mathcal{H}_2$, car pour tout $k \in \mathbb{Z}^2$,

$$(k+a)S^t(k+a) \in 2\mathbb{Z}.$$

En conséquence, la fonction $\vartheta_{a,b}$ admet un développement en série de Fourier de la forme

$$\vartheta_{a,b}(\tau) = \sum_T a_T e^{2i\pi \operatorname{Tr}(T\tau)},$$

où T parcourt les matrices symétriques de $M_2(\mathbb{Q})$ telles que $2n^2T$ est demi-entière. Ce développement en série de Fourier est unique, car on voit directement que pour tout $\tau \in \mathcal{H}_2$,

$$(2n^2)^2 a_T = \int_{[0,1]^4} \vartheta_{a,b}(\tau+x) e^{-2i\pi \operatorname{Tr}(T(\tau+x))} dx.$$

En particulier, la fonction $\vartheta_{a,b}$ est nulle si et seulement si son développement en série de Fourier est nul, c'est pourquoi nous allons calculer celui-ci, qui est presque directement donné par la définition de $\vartheta_{a,b}$ comme série (définition-proposition III.5.22). Notons, pour $a = (a_1, a_2)$ et $k = (k_1, k_2)$:

$$T_{a,k} = \begin{pmatrix} (k_1+a_1)^2 & (k_1+a_1)(k_2+a_2) \\ (k_1+a_1)(k_2+a_2) & (k_2+a_2)^2 \end{pmatrix},$$

de sorte que

$$\vartheta_{a,b}(\tau) = e^{i\pi a^t b} \sum_{k \in \mathbb{Z}^2} e^{2i\pi k^t b} e^{i\pi \operatorname{Tr}(T_{a,k}\tau)}.$$

Ceci n'est pas encore exactement le développement de Fourier de $\vartheta_{a,b}$, en effet il faut regrouper les $T_{a,k}$ donnant le même T. On voit tout de suite que

$$T_{a,k} = T_{a',k'} \iff (k+a) = \pm(k'+a').$$

Si $2a \notin \mathbb{Z}^2$, la fonction $k \to T_{a,k}$ est bijective, et l'expression ci-dessus est bien le développement de Fourier de ϑ, en particulier cette fonction est non nulle.

Sinon, $2a = A \in \mathbb{Z}^2$, et alors pour tous $k, k' \in \mathbb{Z}^2$, on a $(k+a) = \pm(k'+a)$ si et seulement si $k = k'$ ou $k + k' = A$, de sorte que

$$2\vartheta_{a,b}(\tau) = e^{i\pi a^t b} \sum_T \sum_{\substack{k,k' \in \mathbb{Z}^2 \\ T_{k,a} = T_{k',a} = T}} (e^{2i\pi k^t b} + e^{2i\pi(-A-k)^t b}) e^{i\pi \operatorname{Tr}(T\tau)}.$$

Alors, ce développement de Fourier est nul si et seulement si, pour tout $k \in \mathbb{Z}^2$:

$$e^{2i\pi(2k+A)^t b} = -1,$$

c'est-à-dire si et seulement si $b \in (1/2)\mathbb{Z}$ et $(-1)^{4a^t b} = -1$. On retrouve donc exactement la négation de la condition $(*)$.

(c) Soient (a,b) et (a',b') dans $(1/n)\mathbb{Z}^4$ dont les multiples par n vérifient la condition $(*)$. Démontrer le résultat voulu revient à démontrer que $\vartheta_{a,b}^{8n}$ et $\vartheta_{a',b'}^{8n}$ n'ont pas le même diviseur des zéros réduit à moins que $(a,b) = \pm(a',b') \mod \mathbb{Z}^4$. Supposons donc que leur diviseur de zéros réduit est le même. Alors, la fonction

$$\frac{\vartheta_{a,b}^{8n}}{\vartheta_{a',b'}^{8n}}$$

induit une fonction méromorphe sur $A_2(n)^S_{\mathbb{C}}$, car ces deux fonctions sont des formes modulaires de Siegel de degré 2, de niveau n et de même poids $4n$. Par hypothèse, cette fonction méromorphe

a son diviseur des pôles et son diviseur des zéros dans $A_2(n)^S_{\mathbb{C}}$ contenus dans $(D_{n,na,nb})_{\mathbb{C}}$, elle est donc constante. Ainsi, il existe $\lambda \in \mathbb{C}^*$ tel que

$$\vartheta_{a,b} = \lambda \vartheta_{a',b'}.$$

Nous allons comparer les développements de Fourier pour établir le résultat. Tout d'abord, comme $\vartheta_{a,b}$ ne dépend (à une racine de l'unité près) que de $(a,b) \mod \mathbb{Z}^4$ (preuve de la définition-proposition III.5.22), on va supposer que chacun des coefficients de (a,b) et (a',b') appartient à $[-1/2, 1/2[$. Nous allons supposer pour simplifier que ni a ni a' n'appartiennent à $(1/2)\mathbb{Z}^2$: sans cette hypothèse, la preuve est un peu plus calculatoire mais suit les même principes. Alors, avec les notations de la preuve du (b), pour tous $k, k' \in \mathbb{Z}^2$, $T_{k,a} = T_{k',a'}$ impose que $k = k'$ et $a = a'$ (cas 1) ou $k = -k'$ et $a = -a'$ (cas 2) par hypothèse sur a et a'. Dans le cas 1, on a donc pour tout $k \in \mathbb{Z}^2$, par identification des coefficients de Fourier :

$$e^{i\pi a^t b} e^{2i\pi k^t b} = \lambda e^{i\pi a'^t b'} e^{2i\pi k^t b'},$$

donc $k^t(b - b')$ est constant modulo $\mathbb{Z}$ en tant que fonction de de $k \in \mathbb{Z}^2$. Ceci implique que $b = b'$ modulo $\mathbb{Z}^2$, donc $b = b'$ par hypothèse sur les domaines de b et b'. Dans le cas 2, le même raisonnement donne $b = -b'$. On a donc prouvé que si $\vartheta_{a,b}$ et $\vartheta_{a',b'}$ sont proportionnelles, alors $(a,b) = \pm(a',b') \mod \mathbb{Z}^4$, et la réciproque est immédiate.

□

Remarque III.5.9. La preuve du (c) est peu intuitive, mais on peut voir ce résultat d'une autre manière qui se trouve être plus difficile à prouver rigoureusement. Soit un produit de courbes elliptiques A muni d'un fibré ample symétrique L induisant une polarisation principale, et d'une structure symplectique de niveau n. Alors, si le point de n-torsion (mais pas de 2-torsion) de coordonnées (a,b) appartient au diviseur $\Theta_{A,L}$, il est possible de changer la structure symplectique (sans changer de L) pour faire en sorte que le « nouveau » point de coordonnées (a,b) n'appartienne plus à $\Theta_{A,L}$, c'est-à-dire que la fonction $\vartheta_{a,b}$ évaluée en un point $\tau \in \mathcal{H}_2$ associé soit non nulle.

Nous allons maintenant donner les propriétés utiles de ces diviseurs afin de pouvoir appliquer et interpréter un théorème de Runge avec ceux-ci.

Définition-Proposition III.5.29. Soit $n \in \mathbb{N}_{>0}$ pair. On reprend les notations de la définition-proposition III.5.28 et la définition III.5.13.

(a) Pour tout $(a,b) \in (\mathbb{Z}/n\mathbb{Z})^{2g}$ vérifiant la condition $(*)$, le diviseur $(D_{n,a,b})_{\mathbb{C}}$ est ample.

(b) Pour $n = 2$, les dix diviseurs $(D_{2,a,b})_{\mathbb{C}}$ sont disjoints deux à deux en-dehors du bord $\partial A_2(n)_{\mathbb{C}} := A_2(n)^S_{\mathbb{C}} \backslash A_2(n)_{\mathbb{C}}$ de $A_2(n)_{\mathbb{C}}$, et leur réunion est le lieu des modules des produits de courbes elliptiques munis d'une base symplectique quelconque.

(c) Pour (A, λ, α_n) une surface abélienne complexe principalement polarisée avec structure symplectique de niveau n :

- Si (A, λ) est un produit de courbes elliptiques, alors le module de (A, λ, α_n) appartient à exactement $n^2 - 3$ diviseurs $(D_{n,a,b})_{\mathbb{C}}$.

- Sinon, le point (A, λ, α_n) appartient à au plus $(\sqrt{2}/2)n^2$ diviseurs $(D_{n,a,b})_{\mathbb{C}}$.

(d) Pour tous $(a,b) \in (\mathbb{Z}/n\mathbb{Z})^4$ vérifiant la condition $(*)$, le diviseur $(D_{n,a,b})_{\mathbb{C}}$ est la fibre géométrique complexe d'un certain diviseur de Weil effectif réduit $D_{n,a,b}$ de $\mathcal{A}_2(n)$, tel que le module d'un triplet (A, λ, α_n) (défini sur un corps k de caractéristique ne divisant pas n) appartient à $D_{n,a,b}(k)$ si et seulement si le point de $A[n](k)$ de coordonnées (a,b) pour α_n appartient au diviseur $\Theta_{A,\lambda,\alpha_n}$ (définition-proposition III.5.21).

Démonstration. Tout d'abord, le diviseur $(D_{n,a,b})_{\mathbb{C}}$ est le diviseur réduit des zéros d'une forme modulaire de poids $4n$, c'est-à-dire d'une section du fibré $L^{\otimes 4n}$ sur $A_2(n)^S_{\mathbb{C}}$ (définition-proposition III.5.10 (d)). Celui-ci est ample (proposition III.5.8), donc $(D_{n,a,b})_{\mathbb{C}}$ est ample.

Ensuite, d'après la proposition III.5.23, l'image de $\tau \in \mathcal{H}_2$ dans $A_2(n)^S_{\mathbb{C}}$ appartient à $(D_{n,a,b})_{\mathbb{C}}$ si et seulement si le point $\alpha^{-1}_{\tau,n}(a,b)$ de $A_\tau[n]$ appartient à Θ_{A_τ,L_τ}. Alors, si A_τ est une jacobienne, le diviseur Θ_{A_τ,L_τ} contient exactement 6 points de 2-torsion (proposition III.5.27), tous donnés

par les (a,b) ne vérifiant pas la condition $(*)$. Autrement dit, aucun des dix diviseurs $(D_{n,a,b})_{\mathbb{C}}$ (où (a,b) vérifie la condition $(*)$ et est de 2-torsion) ne passe par l'image de τ. Sinon, la surface abélienne A_τ est un produit de courbes elliptiques (proposition III.5.26), et le diviseur Θ_{A_τ,L_τ} contient exactement 7 points de 2-torsion, donc un seul qui est donné par un (a,b) vérifiant la condition $(*)$ (proposition III.5.25). Autrement dit, un et un seul des dix diviseurs $(D_{n,a,b})_{\mathbb{C}}$ (où (a,b) est de 2-torsion) passe par A_τ, ce qui prouve le (b). Les mêmes arguments pour n général, le fait que $D_{n,a,b} = D_{n,-a,-b}$ et les propositions III.5.25 et III.5.27 prouvent le (c).

Enfin, nous allons donner une construction algébrique des $D_{n,a,b}$ pour prouver le (d), tirée de la remarque I.5.2 de [FC90]. Soit $\mathcal{A}$ un schéma abélien sur S (avec morphisme structural $\pi : \mathcal{A} \to S$), et $\mathcal{L}$ un faisceau inversible sur $\mathcal{A}$ relativement ample sur S et symétrique, induisant une polarisation principale sur $\mathcal{A}$. Si $s : S \to \mathcal{A}$ est une section de $\mathcal{A}$ sur S, l'évaluation en s induit un morphisme de $\mathcal{O}_S$-modules entre $\pi_*\mathcal{L}$ et $s^*\mathcal{L}$. Supposons maintenant que s est de n-torsion dans $\mathcal{A}$. Alors, pour $e : S \to A$ la section nulle, le fibré $(s^*\mathcal{L})^{\otimes 2n}$ est isomorphe au fibré $(e^*\mathcal{L})^{\otimes 2n}$, c'est-à-dire trivial. Par ailleurs, si on note $\omega_{\mathcal{A}/S}$ le fibré inversible sur S obtenu comme déterminant du fibré des formes différentielles invariantes par translation sur $\mathcal{A}$, on a $8 \cdot \pi_*\mathcal{L} = -4\omega_{\mathcal{A}/S}$ dans $\mathrm{Pic}(S)$ d'après le théorème I.5.1 et la remarque I.5.2 de [FC90]. En conséquence, l'évaluation en s donnée précédemment définit, après choix d'isomorphisme entre $(e^*\mathcal{L})^{\otimes 2n}$ et $\mathcal{O}_S$, et passage à la puissance $8n$-ième, une section de $\omega_{\mathcal{A}/S}^{\otimes 4n}$ sur S. En appliquant ce résultat au schéma abélien universel $\mathcal{X}_2(n)$ sur $\mathcal{A}_2(n)$, à chaque $(a,b) \in (\mathbb{Z}/n\mathbb{Z})^4$, la section définie par le point de coordonnées (a,b) pour la n-structure sur $\mathcal{A}_2(n)$ induit donc une section globale $s_{a,b}$ de $\omega_{\mathcal{X}_2(n)/\mathcal{A}_2(n)}^{\otimes 4n}$, dont on note $D_{n,a,b}$ le diviseur des zéros réduit de cette section. Maintenant, soit (A, L_A, α_n) un triplet avec A une variété abélienne définie sur un corps k de caractéristique ne divisant pas n, L_A un fibré ample symétrique sur A induisant une polarisation principale, et α_n une n-structure symplectique sur A. Par construction le point de $\mathcal{A}_2(n)$ associé à ce triplet appartient à $D_{n,a,b}$ si et seulement si l'unique section non nulle de L_A (à constante près) s'annule en le point $\alpha_n^{-1}(a,b)$ de $A[n](\overline{k})$, donc si et seulement si ce point appartient au diviseur Θ_{A,L_A}.

En particulier, sur le corps $\mathbb{C}$, le triplet (A, L_A, α_n) appartient à $D_{n,a,b}(\mathbb{C})$ si et seulement si le point de coordonnées (a,b) appartient à Θ_{A,L_A}. Ainsi, pour tout $\tau \in \mathcal{H}_2$, le triplet $(A_\tau, L_\tau, \alpha_{\tau,n})$ appartient à $D_{n,a,b}(\mathbb{C})$ si et seulement si $\alpha_{\tau,n}^{-1}(a,b)$ appartient à Θ_{A_τ,L_τ}, donc si et seulement si l'image de τ dans $A_2(n)_{\mathbb{C}}^S$ appartient à $(D_{n,a,b})_{\mathbb{C}}$ d'après la proposition III.5.23 et la définition de $(D_{n,a,b})_{\mathbb{C}}$. Ceci prouve bien, comme les notations le présageaient, que la fibre géométrique complexe de $D_{n,a,b}$ est exactement $(D_{n,a,b})_{\mathbb{C}}$.

□

Remarque III.5.10. On pourrait voir d'une autre manière la version complexe du (d), en nous basant sur la définition-proposition III.5.4. Pour cela, on définit un certain fibré en droites $\mathcal{L}$ sur $\mathcal{X}_2(n)_{\mathbb{C}}$, qui s'avère être, au-dessus de chaque image de $\tau \in \mathcal{H}_2$ dans $A_2(n)_{\mathbb{C}}$, isomorphe au fibré L_τ, et défini comme un quotient de $\mathcal{H}_2 \times \mathbb{C}^2 \times \mathbb{C}$ par l'action de $\Gamma_2(n) \ltimes \mathbb{Z}^4$ tel que les sections globales de $\mathcal{L}^{\otimes 8n}$ sur $A_2(n)_{\mathbb{C}}$ correspondent canoniquement aux fonctions holomorphes sur $\mathcal{H}_2 \times \mathbb{C}^2$ vérifient une certaine formule de transformation par $\Gamma_2(n) \ltimes \mathbb{Z}^4$, qui se trouve être celle que vérifie la fonction $\Theta(z,\tau)^{\otimes 8n}$ ([Mum87], section II.5 combinée avec la formule (III.7)). Ainsi, l'évaluation le long de la section nulle nous donne exactement la fonction $\Theta(0,\tau)^{8n}$, donc le diviseur $(D_{n,0,0})_{\mathbb{C}}$, et on retrouve de même les autres diviseurs avec les autres sections de n-torsion de $\mathcal{X}_2(n)_{\mathbb{C}}$.

Nous pouvons maintenant prouver deux théorèmes de Runge tubulaire pour les diviseurs thêta. Le cas $n \geq 4$ sera précédé par le cas $n = 2$, où la situation est particulièrement intuitive, car le « bord » (lieu par rapport auquel on définit l'intégralité) auquel on applique le théorème de Runge tubulaire a une bonne interprétation modulaire comme lieu de dégénérescence de surfaces abéliennes. En effet, il se compose du « bord fort », lieu des variétés semi-abéliennes avec partie torique non nulle (pour l'intégralité parfaite) et du « bord faible », lieu des surfaces abéliennes principalement polarisées non simples, dont les composantes constituent les diviseurs effectifs définissant l'intégralité usuelle pour le théorème.

C'est une situation analogue au cas des courbes modulaires, où le bord correspond au lieu des courbes elliptiques dégénérées, c'est-à-dire aux pointes, et la distinction entre bord fort et bord

faible n'y a pas de sens.

Théorème III.9 (Runge pour les produits de courbes elliptiques sur $\mathcal{A}_2(2)^S$)**.**

Pour un voisinage ouvert U du bord $\partial A_2(2)_{\mathbb{C}} := A_2(2)^S_{\mathbb{C}} \backslash A_2(2)_{\mathbb{C}}$ pour la topologie complexe usuelle, soit $\mathcal{E}(U)$ l'ensemble des points $P \in A_2(2)(\overline{\mathbb{Q}})$ représentant le triplet (A, λ, α_2) (à isomorphisme à extension des scalaires près), avec A une surface abélienne, λ une polarisation principale sur A et α_2 une structure symplectique de niveau 2 tel que :

- *Le point P est à valeurs dans une certaine extension finie L de $\mathbb{Q}$.*
- *La variété abélienne A a potentiellement bonne réduction en toute place finie (c'est-à-dire que P ne rencontre pas le bord fort), et quelle que soit l'extension des scalaires de $A_2(2)(L)$ à $A_2(2)_{\mathbb{C}}$, l'image du point P n'appartient pas à l'ouvert U.*
- *Le nombre s_L de places v de L telles que*
  - *v est archimédienne*
  - *v divise 2*
  - *la réduction modulo v de (A, λ) est à extension des scalaires près un produit de courbes elliptiques (c'est-à-dire appartient au support d'un des dix diviseurs $D_{2,a,b}$: condition de bord faible)*

vérifie la condition de Runge

$$s_L < 10.$$

Alors, cet ensemble $\mathcal{E}(U)$ est fini pour tout ouvert U contenant $\partial A_2(2)_{\mathbb{C}}$.

Théorème III.10 (Runge pour les diviseurs thêta)**.**

Soit $n \geq 2$ un entier pair. Pour un voisinage ouvert U du bord $\partial A_2(n)_{\mathbb{C}} := A_2(n)^S_{\mathbb{C}} \backslash A_2(n)_{\mathbb{C}}$ (pour la topologie complexe), soit $\mathcal{E}(U)$ l'ensemble des points $P \in A_2(n)(\overline{\mathbb{Q}})$ représentant le triplet (A, λ, α_n) (avec A une surface abélienne, λ une polarisation principale sur A et α_n une structure de niveau n) tels que :

- *Le point P est défini sur une certaine extension finie L de $\mathbb{Q}(\zeta_n)$.*
- *La variété abélienne A a potentiellement bonne réduction en toute place finie, et quelle que soit l'extension des scalaires de $A_2(n)(L)$ à $A_2(n)_{\mathbb{C}}$, l'image du point P n'appartient pas à l'ouvert U.*
- *Le nombre s_L de places v de L telles que*
  - *v est archimédienne*
  - *v divise n*
  - *la réduction modulo v de (A, λ) admet non trivialement un point de n-torsion dans son diviseur thêta (c'est-à-dire que la réduction de (A, λ, α_n) appartient au support d'un des $D_{n,a,b}$: condition de bord faible)*

vérifie

$$(n^2 - 3)s_L < n^4/2 + 2$$

(condition de Runge).

Alors, cet ensemble $\mathcal{E}(U)$ est fini pour tout ouvert U contenant $\partial A_2(n)_{\mathbb{C}}$.

Preuve du théorème III.10. Il s'agit d'une application du théorème de Runge tubulaire (théorème III.8) à $X = \mathcal{A}_2(n)^S$ et aux diviseurs $D_{n,a,b}$ définis précédemment, vérifions-en les hypothèses. Avec les notations de ce théorème, le corps de base est $K = \mathbb{Q}(\zeta_n)$, et S_0 est l'ensemble des places de K qui sont archimédiennes ou divisent n (définition-proposition III.5.14 et définition-proposition III.5.29 (d)). Ensuite, les $D_{n,a,b}$ sont bien des diviseurs de Weil effectifs sur $\mathcal{A}_2(n)^S$ au nombre de $n^4/2 + 2$ donc de Cartier effectif quitte à multiplier par un entier (proposition III.5.16 (b)), et les $(D_{n,a,b})_K$ sont amples (définition-proposition III.5.29 (a)). De plus, l'union de leurs supports dans $A_2(n)_{\mathbb{C}}$ est le lieu des modules de variétés abéliennes (A, λ, α_n) dont le diviseur thêta contient un point de n-torsion *non trivial*, c'est-à-dire ne provenant pas des six (a, b) ne vérifiant pas la condition $(*)$ ((définition-propositions III.5.28) (c) et III.5.29 (b)). Hors de $\partial\mathcal{A}_2(n)_{\mathbb{C}}$, on sait de plus que l'intersection de $n^2 - 2$ de ces diviseurs est vide (définition-proposition III.5.29 (d)), et on choisit donc Y dans ce théorème comme le bord $\partial\mathcal{A}_2(n) := \mathcal{A}_2(n)^S \backslash \mathcal{A}_2(n)$ (de sorte que $m = n^2 - 3$). Ainsi, la condition d'être $X\backslash Y$-entier en une place finie v non au-dessus de

S_0 signifie que la variété abélienne sous-jacente a bonne réduction en cette place, reste donc à interpréter ce que veut dire être parfaitement $X\backslash Y$-entier. Les invariants d'Igusa, qui sont des formes modulaires sur $\mathcal{A}_2(1)$ donc sur $\mathcal{A}_2(n)$, caractérisent le bord $\partial\mathcal{A}_2(n)$ comme ensemble de leurs zéros communs [Igu60] et on les choisit donc comme générateurs de l'idéal de définition de Y, de sorte que si la variété abélienne A a bonne réduction en une place finie v au-dessus de S_0, son module dans $\mathcal{A}_2(n)$ est $X\backslash Y$-entier en v pour ces générateurs. Enfin, on assouplit la condition de $X\backslash Y$-intégralité parfaite en une quasi-$X\backslash Y$-intégralité (définition III.4.5) en choisissant une M_K-constante nulle sur les places finies, mais telle que pour toute place archimédienne v sur $\overline{K}$ et tout point $P \in A_2(n)(\overline{K})$, si $P_v \notin U$ pour un plongement complexe de P associé à v, le point P est quasi-$X\backslash Y$-entier en v pour notre M_K-constante. On peut donc appliquer le théorème III.8 (avec l'ajout de la remarque III.4.6) avec toutes ces données pour obtenir le résultat.

Ainsi, sous les hypothèses de l'énoncé (la condition de Runge tubulaire se traduit en la condition $(n^2-3)s_L < n^4/2+2$ d'après les comptages précédents), l'ensemble $\mathcal{E}(U)$ est bien fini. □

Preuve du théorème III.9. Ce théorème est un cas particulier du théorème précédent, nous allons donc simplement montrer d'où vient la condition de Runge dans ce cas. Nous avons ici dix diviseurs de Weil sur $\mathcal{A}_2(1)^S$, dont nous savons par la définition-proposition III.5.29 (b) qu'ils sont disjoints hors du bord (d'où $m=1$ et $r_{\mathbb{Q}}=10$ avec les notations du théorème III.8). De plus, ils caractérisent à extension des scalaires près le fait d'être une courbe elliptique. On peut donc appliquer le théorème III.8 comme pour la preuve ci-dessus, le seul changement étant cette interprétation en terme de produits de courbes elliptiques, disponible pour $n=2$ mais pas pour $n \geq 4$. □

Remarque III.5.11. Il serait plaisant de pouvoir obtenir un théorème de Runge pour les composantes irréductibles du lieux de produits de courbes elliptiques avec tout niveau n. Malheureusement, cela ne semble pas accessible pour le moment sauf pour $n=2$, car la question de la grosseur de ces diviseurs reste ouverte, contrairement à celle des diviseurs thêta (qui sont même amples). Ensuite, une manière agréable de voir le théorème III.9 (qui pourrait se formuler aussi pour sa généralisation) est de le comprendre comme un théorème de concentration des points entiers près du bord (remarque III.4.6).

Pour que ce théorème soit non vide, il faut s'assurer que le nombre de places minimal s_L d'un des points ci-dessus peut bien vérifier la condition de Runge. Le lemme suivant l'assure, et donne une idée de la marge que la condition de Runge autorise pour les extensions L et les places comprises dans le comptage de s_L. La preuve de ce résultat est très élémentaire, mais nous la donnons par manque de références à ce sujet dans la littérature.

Lemme III.5.30. *Pour tout entier $n \geq 4$ pair, le nombre de places de $\mathbb{Q}(\zeta_n)$ archimédiennes ou divisant n est inférieur à $n/2$.*

Exemple III.5.1. Pour $n=4$, le plus petit cas possible, le nombre de places archimédiennes (ou au-dessus de 2) de $\mathbb{Q}(\zeta_4)$ est égal à 2, et le rapport $(4^4/2+8)/(4^2-3)$ est égal à 10, on peut donc choisir 7 places de plus pour les points définis sur $\mathbb{Q}(\zeta_4)$ et encore assurer la finitude.

Démonstration. Notons φ l'indicatrice d'Euler, et $\prod_{p|n} f(p)$ et $\sum_{p|n} f(p)$ le produit et la somme sur les nombres premiers divisant n des valeurs d'une certaine fonction f en ces nombres premiers. Par exemple, $n\prod_{p|n}(1-p^{-1}) = \varphi(n)$. Notons P_n le nombre de places de $\mathbb{Q}(\zeta_n)$ archimédiennes ou au-dessus de n

Pour p premier divisant n, on note $n_p = n/p^\alpha$ avec α la multiplicité de p dans n (de sorte que p ne divise pas n_p) et $o_{n,p}$ l'ordre de p dans $(\mathbb{Z}/n_p\mathbb{Z})^*$.

Comme $n \geq 3$, aucun plongement de $\mathbb{Q}(\zeta_n)$ n'est réel et il y a donc $\varphi(n)/2$ places infinies de $\mathbb{Q}(\zeta_n)$.

Pour tout nombre premier p divisant n, comptons maintenant le nombre de places de $\mathbb{Q}(\zeta_n)$ au-dessus de p. Par des considérations classiques de théorie des nombres, si Φ_n est le n-ième polynôme cyclotomique et $\overline{\Phi_n}$ sa réduction modulo p, le nombre de places de $\mathbb{Q}(\zeta_n)$ au-dessus de p est égal au nombre de facteurs irréductibles distincts de $\overline{\Phi_n}$ dans $\mathbb{F}_p[X]$. Ceux-ci sont donnés par les orbites du Frobenius parmi les racines de $\overline{\Phi_n}$ dans $\overline{\mathbb{F}_p}$, et correspondent donc aux orbites de

la mise à la puissance p dans $(\mathbb{Z}/n_p\mathbb{Z})^*$. Il y a donc exactement $\varphi(n_p)/o_{n,p}$ facteurs irréductibles distincts de $\overline{\Phi_n}$, d'où la formule exacte

$$P_n = \frac{\varphi(n)}{2} + \sum_{p|n} \frac{\varphi(n_p)}{o_{n,p}}. \tag{III.8}$$

Dans le cas où $n = p^\alpha, \alpha \geq 2$ et p premier (pouvant être égal à 2),

$$P_n = p^{\alpha-1}(p-1)/2 + 1 \leq p^\alpha/2 = \frac{p^\alpha + 1 - p^{\alpha-1}}{2} \leq \frac{p^\alpha}{2}.$$

Ensuite, remarquons que si n' est un diviseur de n avec les mêmes facteurs premiers, $\varphi(n)/n = \varphi(n')/n'$ et que pour $p|n$, $o_{n,p} \geq o_{n',p}$ donc $P_{n'}/n' \geq P_n/n$. Comme $P_4 = 2$, il reste donc à démontrer le résultat pour n sans facteur carré et différent de 2, ce qu'on suppose dorénavant.

Comme $o_{n,p}$ est un entier non nul tel que $p^{o_{n,p}} = 1 \mod n_p$, il est au moins égal à $\ln n_p / \ln p$, en particulier il est au moins égal à 2 sauf si $n_p < p$, ce qui est possible seulement pour le plus grand diviseur premier de n.

Si on note q le plus grand diviseur premier de n, on a donc l'inégalité

$$\frac{P_n}{n} \leq \frac{1}{2}\prod_{p|n}(1-p^{-1}) + \frac{1}{2}\sum_{\substack{p|n\\p<q}} \frac{1}{p} \prod_{\substack{p'|n\\p'\neq p}} (1-p'^{-1}) + \frac{1}{qo_{n,q}} \prod_{\substack{p'|n\\p'\neq q}} (1-p'^{-1}). \tag{III.9}$$

En indexant dans l'ordre croissante les facteurs premiers $p_1 = 2, \cdots, p_r = q$ de n, on remarque que

$$\prod_{p|n}(1-p^{-1}) + \frac{1}{p_1}\prod_{i=2}^{r}(1-p_i^{-1}) = \prod_{i=2}^{r}(1-p_i^{-1})$$

On peut ainsi éliminer un par un les termes correspondant aux facteurs premiers (sauf les deux plus grands, notés p et q) dans la deuxième somme pour obtenir

$$\frac{P_n}{n} \leq \frac{1}{2}(1-p^{-1})(1-q^{-1}) + \frac{1}{2p}(1-q^{-1}) + \frac{1}{2o_{pq,q}}(1-p^{-1})$$

(car $o_{n,q} \geq o_{pq,q}$).

Dans le cas où $o_{pq,q} \geq 2$, on peut réutiliser une dernière fois l'astuce d'élimination pour obtenir la borne 1/2. Sinon, $o_{pq,q} = 1$ et alors $q = 1 \mod p$. Mais en remontant les inégalités précédentes, on a $o_{pq,p} \geq 3$ car sinon $p^2 = 1 \mod q$, donc q divise $p-1$ ou $p+1$, forcément $p+1$ par comparaison, et alors $q = p+1$ ce qui impose $p = 2$, $q = 3$ et $n = 6$, mais $P_6 = 3 \leq 6/2$. On peut donc supposer que $o_{pq,q} \geq 3$, et alors

$$\frac{P_n}{n} \leq \frac{1}{6pq}(3(p-1)(q-1) + 2(q-1) + 6(p-1)) = \frac{1}{6pq}(3pq + 3p - q - 8)$$

ce qui est bien inférieur à 1/2 à moins que $q < 3p - 8$, et alors $q = 1 + 2p$. Mais alors, $o_{pq,p} \geq 4$ car sinon $1 + 2p$ divise $1 + p + p^2$ ce qui est impossible. On peut donc reprendre l'inégalité (III.9) avec $o_{n,p} \geq 4$, et refaire l'élimination des autres termes pour obtenir

$$\frac{P_n}{n} \leq \frac{1}{4pq}2(p-1)(q-1) + (q-1) + 4(p-1) = \frac{1}{4pq}(2pq + 2p - q - 3) \leq 1/2$$

car on est encore dans le cas où $q = 1 + 2p$. □

Bibliographie

[ABMMOG14] E. Artal Bartolo, J. Martín-Morales, and J. Ortigas-Galindo. Cartier and Weil divisors on varieties with quotient singularities. *Int. Journ. Math.*, 25(11), 2014.

[Akb97] A. Akbary. Non-vanishing of modular L-functions with large level. PhD Thesis, http ://www.cs.uleth.ca/∼akbary/publications.html, 1997.

[AL70] A. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185 :134–160, 1970.

[Ale02] V. Alexeev. Complete moduli in the presence of semiabelian group action. *Ann. of Math. (2)*, 155(3) :611–708, 2002.

[AM94] M. Atiyah and I. Macdonald. *Introduction to commutative algebra*. Westview Press, 1994.

[AMRT10] Ash, Mumford, Rapoport, and Tai. *Smooth compactifications of locally symmetric varieties*. Cambridge University Press, 2nd edition, 2010.

[Apo90] T. Apostol. *Modular functions and Dirichlet series in number theory*, volume 41 of *Graduate Texts in Mathematics*. Springer-Verlag, 2nd edition, 1990.

[Bai58] W. Baily. Satake's compactification of V_n. *Amer. J. Math.*, 80 :348–364, 1958.

[BD97] M. Bertolini and H. Darmon. A Rigid Analytic Gross-Zagier Formula and Arithmetic Applications. *Ann. of Math.*, 146(1) :111–147, 1997.

[BD10] N. Billerey and L. Dieulefait. Solving Fermat-type equations $x^5+y^5 = dz^p$. *Math. Comp.*, 79 :535–544, 2010.

[BEN10] M. Bennett, J. Ellenberg, and N. Ng. The diophantine equation $A^4 + 2^\delta B^2 = C^n$. *Int. J. Number Theory*, 6 :311–338, 2010.

[BG00] J. Boxall and D. Grant. Examples of torsion points on genus two curves. *Trans. Amer. Math. Soc.*, 352(10) :4533–4555, 2000.

[BG06] E. Bombieri and W. Gubler. *Heights in diophantine geometry*. Cambridge University Press, 2006.

[BL04] C. Birkenhake and H. Lange. *Complex abelian varieties*. Springer-Verlag, Berlin, 2nd edition, 2004.

[BLR90] S. Bosch, W. Lütkebohmert, and M. Raynaud. *Néron models*. Springer-Verlag, 1990.

[Bom83] E. Bombieri. On Weil's "Théorème de décomposition". *Amer. J. Math.*, 105(2) :295–308, 1983.

[Bor81] A. Borel. Stable real cohomology of arithmetic groups II. In *Manifolds and Lie groups*, pages 21–55. Birkhäuser, 1981.

[BP11a] Y. Bilu and P. Parent. Serre's Uniformity Problem in the Split Cartan case. *Ann. of Math. (2)*, 173 :569–584, 2011.

[BP11b] Y. Bilu and P. Parent. Runge's method and modular curves. *Int. Math. Res. Not.*, (9) :1997–2027, 2011.

[BPR13] Y. Bilu, P. Parent, and M. Rebolledo. Rational points on $X_0^+(p^r)$. *Annales de l'Institut Fourier*, 63, 2013.

[BT08] F. Beukers and S. Tengely. An implementation of Runge's method for Diophantine equations. arXiv :1103.5388, 2008.

[Bum96] D. Bump. *Automorphic forms and representations.* Cambridge University Press, 1996.

[BvdGHZ08] Bruinier, van der Geer, Harder, and Zagier. *The 1-2-3 of modular forms.* Universitext. Springer-Verlag, Berlin, 2008.

[Car57] H. Cartan. Plongements projectifs. *Séminaire Henri Cartan*, 10(2) :1–19, 1957.

[CFM13] D. Chen, G. Farkas, and I. Morrison. Effective divisors on moduli spaces of curves and abelian varieties. In *A Celebration of Algebraic Geometry*, pages 131–169. Amer. Math. Soc., 2013.

[Che00] I. Chen. On Relations between Jacobians of Certain Modular Curves. *Journal of Algebra*, 231(1) :414–448, 2000.

[CLZ09] P. Corvaja, A. Levin, and U. Zannier. Integral points on threefolds and other varieties. *Tohoku Mathematical Journal*, 61 :589–601, 2009.

[Coj05] A. C. Cojocaru. On the Surjectivity of the Galois Representations Associated to Non-CM Elliptic Curves. *Canad. Math. Bull.*, 48 :16–31, 2005.

[CS57] H. Cartan and I. Satake. Démonstration du théorème fondamental. *Séminaire Henri Cartan*, 10(2) :1–12, 1957.

[Dar09] H. Darmon. Rational points on curves. In *Arithmetic geometry*, pages 7–53. Amer. Math. Soc., 2009.

[Deb99] O. Debarre. *Tores et variétés abéliennes complexes.* EDP Sciences, 1999.

[DM97] H. Darmon and L. Merel. Winding quotients and some variants of Fermat's Last Theorem. *J. Reine Angew. Math.*, 490 :81–100, 1997.

[DR73] P. Deligne and M. Rapoport. Les schémas de modules de courbes elliptiques. In *Modular functions of one variable II*, pages 143–316. Springer, 1973.

[DS05] F. Diamond and J. Shurman. *A First Course in Modular Forms.* Graduate Texts in Mathematics. 2005.

[dSE00] B. de Smit and B. Edixhoven. Sur un résultat d'Imin Chen. *Mat. Res. Lett.*, 7 :147–153, 2000.

[Duk95] W. Duke. The critical order of vanishing of automorphic L-functions with large level. *Invent.Math.*, 119 :165–174, 1995.

[Edi84] B. Edixhoven. Rational torsion points on elliptic curves over number fields. *Séminaire Bourbaki*, pages 209–227, 1983-1984.

[Elk04] N. Elkies. On Elliptic K-curves. In *Modular Curves and Abelian Varieties*, volume 224 of *Prog. Math.*, pages 81–91. Birkhäuser, 2004.

[Ell04] J. Ellenberg. Galois Representations Attached to $\mathbb{Q}$-Curves and the Generalized Fermat Equation $A^4 + B^2 = C^p$. *Amer. J. Math.*, pages 763–787, 2004.

[FC90] G. Faltings and C.-L. Chai. *Degeneration of abelian varieties.* Springer-Verlag, 1990.

[GL98] J. González and J.-C. Lario. Rational and Elliptic parametrizations of $\mathbb{Q}$-curves. *J. Number Theory*, 72(1) :13–31, 1998.

[GR14] E. Gaudron and G. Rémond. Théorème des périodes et degrés minimaux d'isogénies. *Comment. Math. Helv.*, 2014.

[Har77] R. Hartshorne. *Algebraic Geometry.* Springer-Verlag, 1977.

[Has97] Y. Hasegawa. $\mathbb{Q}$-curves over quadratic fields. *Manuscripta Math.*, 94 :347–364, 1997.

[HM90] J. Harris and I. Morrison. Slopes of effective divisors on the moduli space of stable curves. *Invent. Math.*, 99(2) :321–355, 1990.

[HS83] D. Hilliker and E. Straus. Determination of bounds for the solutions to those binary Diophantine equations that satisfy the hypotheses of Runge's theorem. *Trans. Amer. Math. Soc.*, 280(2) :637–657, 1983.

[HS00] M. Hindry and J. Silverman. *Diophantine Geometry : an Introduction*. Springer, 2000.

[Hul00] K. Hulek. Nef divisors on moduli spaces of abelian varieties. In *Complex analysis and algebraic geometry*, pages 255–274. de Gruyter, Berlin, 2000.

[HW00] J. Hoffman and S. Weintraub. The Siegel Modular Variety of Degree Two and Level Three. *Trans. Amer. Math. Soc.*, 353(3) :3267–3305, 2000.

[HW08] G. Hardy and E. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 6th edition, 2008.

[Igu60] J.-i. Igusa. Arithmetic variety of moduli for genus two. *Ann. of Math. (2)*, 72 :612–649, 1960.

[Igu67a] J.-I. Igusa. A Desingularization Problem in the Theory of Siegel Modular Functions. *Math. Ann.*, pages 228–260, 1967.

[Igu67b] J.-i. Igusa. Modular forms and projective invariants. *Amer. J. Math.*, 89 :817–855, 1967.

[Igu72] J.-I. Igusa. *Theta Functions*. Springer-Verlag, 1972.

[IK04] H. Iwaniec and E. Kowalski. *Analytic number theory*. Amer. Math. Soc., 2004.

[Kat04] K. Kato. p-adic Hodge theory and values of zeta functions of modular forms. *Astérisque*, (295) :ix, 117–290, 2004.

[KL81] D. Kubert and S. Lang. *Modular units*. Springer-Verlag, 1981.

[KL90] V. Kolyvagin and D. Logachëv. Finiteness of the Shafarevich-Tate group and the group of rational points for some modular abelian varieties. *Leningrad Math. J.*, (1) :1229–1253, 1990.

[Kli90] H. Klingen. *Introductory Lectures on Siegel Modular Forms*. Cambridge University Press, 1990.

[KW09a] C. Khare and J.-P. Wintenberger. Serre's modularity conjecture (I). *Invent. Math.*, 178 :485–504, 2009.

[KW09b] C. Khare and J.-P. Wintenberger. Serre's modularity conjecture (II). *Invent. Math.*, 178 :505–586, 2009.

[Lan02] S. Lang. *Algebra*. Graduate Texts in Mathematics. Springer-Verlag, 2002.

[Laz04] R. Lazarsfeld. *Positivity in algebraic geometry. I*. Springer-Verlag, 2004.

[Lev08] A. Levin. Variations on a theme of Runge : effective determination of integral points on certain varieties. *J. Théor. Nombres Bordeaux*, pages 385–417, 2008.

[LF] S. Le Fourn. Surjectivity of Galois representations associated with quadratic $\mathbb{Q}$-curves. *Math. Ann.* to appear.

[LF15] S. Le Fourn. Nonvanishing of central values of L-functions of newforms in $S_2(\Gamma_0(dp^2))$ twisted by quadratic characters, 2015. arXiv :1506.08723.

[Liu02] Q. Liu. *Algebraic Geometry and Arithmetic Curves*. Oxford University Press, 2002.

[Maz77] B. Mazur. Modular curves and the Eisenstein ideal. *Pub. math. IHES*, 47 :33–186, 1977.

[Maz78] B. Mazur. Rational isogenies of prime degree (with an appendix by D. Goldfeld). *Invent. Math.*, 44(2) :129–162, 1978.

[Mer07] L. Merel. Normalizers of split Cartan subgroups and supersingular elliptic curves. *Diophantine Geometry, CRM Series*, 4 :237–255, 2007.

[Mom84] F. Momose. Rational points on the modular curves $X_{\mathrm{split}}(p)$. *Comp. Math.*, 52(1) :115–137, 1984.

[MS02] F. Momose and M. Shimura. Lifting of supersingular points on $X_0(p^r)$ and lower bound of ramification index. *Nagoya Mathematical Journal*, 165 :159–178, 2002.

[Mum83] D. Mumford. Towards an Enumerative Geometry of the Moduli Space of Curves. In *Arithmetic and Geometry*, volume 36, pages 271–328. Birkhäuser, 1983.

[Mum86] D. Mumford. *Abelian Varieties*. Oxford University Press, 1986.

[Mum87] D. Mumford. *Tata Lectures on Theta I*. Birkhäuser, 1987.

[Mum99] D. Mumford. *The Red Book of Varieties and Schemes*. Lecture notes in mathematics. Springer-Verlag, 1999.

[Nam80] Y. Namikawa. *Toroidal compactification of Siegel spaces*, volume 812 of *Lecture Notes in Mathematics*. Springer, 1980.

[Nér64] A. Néron. Modèles minimaux des variétés abéliennes sur les corps locaux et globaux. *Pub.Math. IHES.*, 21 :128, 1964.

[OU73] F. Oort and K. Ueno. Principally polarized abelian varieties of dimension two or three are Jacobian varieties. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 20 :377–381, 1973.

[Paz13] F. Pazuki. Minoration de la hauteur de Néron-Tate sur les surfaces abéliennes. *Manuscripta Math.*, 142(1-2) :61–99, 2013.

[Pom11] C. Pomerance. Remarks on the Pólya-Vinogradov inequality. In *Integers*, volume 11, pages 531–542, 2011.

[Poo98] B. Poonen. On some diophantine equations. *Acta Arithmetica*, 86, 1998.

[Ran77] R. Rankin. *Modular forms and functions*. Cambridge University Press. 1977.

[Rib04] K. Ribet. Abelian Varieties over $\mathbb{Q}$ and Modular Forms. In *Modular Curves and Abelian Varieties*, pages 241–261. Birkhäuser, 2004.

[Run87] C. Runge. Ueber ganzzahlige lösungen von gleichungen zwischen zwei veränderlichen. *J. Reine Angew. Math.*, 100 :425–435, 1887.

[Sch08] R. Schoof. *Catalan's Conjecture*. Springer-Verlag, 2008.

[Ser72] J.-P. Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15(4) :259–331, 1972.

[Ser97] J.-P. Serre. *Lectures on the Mordell-Weil theorem*. Aspects of mathematics. F. Vieweg, 1997.

[Shi72] T. Shioda. On elliptic modular surfaces. *J. Math. Soc. Japan*, 24(1) :20–59, 1972.

[Sil94] J. Silverman. *Advanced Topics in The Arithmetic of Elliptic Curves*. Graduate texts in mathematics. Springer-Verlag, 1994.

[Sil09] J. Silverman. *The Arithmetic of Elliptic Curves*, volume 106. Springer-Verlag, 2009.

[Spr81] V. Sprindžuk. Reducibility of polynomials and rational points on algebraic curves. In *Seminar on Number Theory, Paris 1979–80*, pages 287–309. Birkhäuser, 1981.

[Suz82] M. Suzuki. *Group Theory : Volume I*. Springer-Verlag, 1982.

[vdG98] G. van der Geer. The Chow ring of the moduli space of abelian threefolds. *J. Algebraic Geom.*, 7(4) :753–770, 1998.

[Wat22] G. Watson. *Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library. 1922.

[Wei92] R. Weissauer. The Picard group of Siegel modular threefolds. *J. Reine Angew. Math.*, 430 :179–211, 1992.