Représentations galoisiennes et groupe de Mumford-Tate associé à une variété abélienne

Davide Lombardo

NNT : 2015SACLS196

THÈSE DE DOCTORAT
DE
L'UNIVERSITÉ PARIS-SACLAY

PRÉPARÉE À
L'UNIVERSITÉ PARIS-SUD XI

ÉCOLE DOCTORALE N° 574
Mathématiques Hadamard

Spécialité : Mathématiques fondamentales

par

M. Davide Lombardo

Représentations galoisiennes et groupe de Mumford-Tate associé à une variété abélienne

(Galois representations and Mumford-Tate groups
attached to abelian varieties)

Thèse présentée et soutenue à Orsay, le 10/12/2015.

Composition du jury :

Daniel	BERTRAND	Université Pierre et Marie Curie	Président
Anna	CADORET	École Polytechnique	Examinatrice
Guy	HENNIART	Université Paris-Sud	Examinateur
Pierre	PARENT	Université de Bordeaux	Examinateur
Nicolas	RATAZZI	Université Paris-Sud	Directeur de thèse
Jean-Pierre	WINTENBERGER	Université de Strasbourg	Rapporteur

Rapporteur externe :

Kenneth A.	RIBET	University of California	Rapporteur

Thèse préparée au
Département de Mathématiques d'Orsay
Laboratoire de Mathématiques (UMR 8628), Bât. 425
Université Paris-Sud 11
91405 Orsay CEDEX

"Many things went on at Unseen University and, regrettably, teaching had to be one of them. The faculty had long ago confronted this fact and had perfected various devices for avoiding it. But this was perfectly all right because, to be fair, so had the students."

Sir Terry Pratchett

Table des matières

Keywords: Galois representations, Mumford-Tate conjecture, elliptic curves, abelian varieties, open image theorems, profinite groups.
2000 Mathematics Subject Classification: 11G10, 11G05, 11F80, 14K15, 14K22.

Remerciements

D'abord et avant tout, mes remerciements vont à Nicolas Ratazzi, qui a dirigé ce travail, pour avoir partagé avec moi ses idées et ses connaissances et pour m'avoir consacré beaucoup de son temps et de son énergie. Au moment de la rédaction de ces remerciements, j'avais échangé avec lui le nombre record de 684 courriels, soit une moyenne de deux messages tous les trois jours au cours des trois dernières années : cela donne une mesure de son immense disponibilité, mais ne suffit pas à montrer sa capacité à me guider, à me stimuler et à me débloquer quand j'en avais besoin. Cette thèse n'aurait pas pu démarrer, ni progresser, ni se conclure sans lui, et je ne sais pas comment lui témoigner suffisamment ma gratitude.

Jean-Pierre Wintenberger et Kenneth Ribet ont rédigé des rapports sur cette thèse, et je les en remercie chaleureusement. Je suis immensément honoré de les avoir eus comme rapporteurs et je suis heureux que le premier ait accepté d'assister à ma soutenance.

Je suis également très heureux que Pierre Parent et Daniel Bertrand aient accepté de faire partie de mon jury de thèse. Il en va bien sûr de même pour Guy Henniart et Anna Cadoret, que je tiens à remercier pour la gentillesse qu'ils m'ont toujours montrée.

Je tiens aussi à exprimer ma gratitude à David Harari : son cours sur la théorie du corps de classes (que j'ai eu la chance de suivre lors de mon M2 à Orsay) s'est révélé être une boite à outils vraiment indispensable pour tout ce que j'ai fait depuis. Je tiens également à le remercier pour avoir accepté à plusieurs reprises de répondre à mes questions sur l'arithmétique des tores.

Je sais gré à Etienne Fouvry pour sa disponibilité et pour m'avoir poussé à participer au groupe de travail de théorie analytique des nombres.

Je suis reconnaissant envers Gaël Rémond et Éric Gaudron pour avoir répondu en plusieurs occasions à mes questions sur leurs travaux sur le théorème d'isogénie – même quand ces questions n'avaient pas de sens ! – toujours avec grande gentillesse et générosité, et pour avoir relevé de nombreuses coquilles dans des textes que je leur avais envoyés.

Cela a été un très grand honneur de pouvoir discuter en personne des thèmes de cette thèse avec Jean-Pierre Serre : sa profonde influence sur le sujet, son immense culture mathématique et en même temps sa disponibilité sont un modèle auquel on peut seulement aspirer. Je le remercie notamment pour une longue discussion à propos de ses travaux qui eut lieu à l'Institut Henri Poincaré, et grâce à laquelle j'ai eu accès aux secrets de la bibliothèque de l'IHP.

Je tiens également à exprimer ma gratitude à Marc Hindry, dont j'admire énormément le talent et qui a toujours été très disponible pour répondre à mes questions, ce qu'il a invariablement fait avec grande clarté. Les discussions que j'ai eues avec lui ont toujours été très enrichissantes et instructives, et sa présence discrète une grande inspiration.

Je remercie Antonella Perucca pour l'intérêt qu'elle a porté au sujet du chapitre 7 de ma thèse, et pour ses remarques pertinentes qui ont beaucoup contribué à améliorer ce chapitre (ou du moins, je l'espère !).
A l'occasion d'un exposé récent, les membres de l'équipe de théorie des nombres de Clermont-Ferrand – surtout Marusia Rebolledo, Nicolas Billerey, Pierre Lezowski et Eric Gaudron – m'ont réservé un accueil chaleureux, pour lequel je les remercie vivement. À cette même occasion j'ai aussi eu la chance de profiter de la compagnie de Samuel Le Fourn, dont j'ai apprécié l'intelligence et l'amitié.

Côté divertissement, le Club Intercontinental de Lecture – Francesco, Denis, Alessandra, Marta et Jinglei (et Gennady, même s'il n'était là qu'en esprit) – a été une compagnie constante et très agréable pendant ces deux dernières années : pour ça, et pour m'avoir souvent donné l'excuse pour une bonne tasse de thé, un grand merci.
Je souhaite saluer ici aussi les doctorants de l'équipe AGA d'Orsay, en particulier Tiago, Lucile, Cong et Yang ; un clin d'oeil aussi à Margaret, Kevin, Victoria, Cagri et Thibault, qui me font sentir chez moi en France (tâche ardue !), et encore un grand merci à Lucile, qui a corrigé mon français.
Je remercie Marco, Fabrizio, Michele et Sara, Daniele, Gennady et Giovanni pour les (Scrocco) Colloquia ; Giovanni pour m'avoir accueilli chez lui quand j'en avais besoin, aussi bien que pour être un excellent joueur de *King of New York* ; Fabrizio pour m'avoir appris que – dans le doute – il faut toujours dériver les produits scalaires ; Marco qui a été mon guide à travers les Marais de la Topologie et pour des jours merveilleux passés à Londres ; Daniele pour ses fréquents courriels, que je ne mérite pas ; Michele et Sara pour leur hospitalité, dont j'ai eu souvent l'occasion de profiter ; Gennady, parce que – même s'il ne répond pas aux emails – c'est un vrai ami. Merci aussi à mes camarades d'adoption, Nicola, Valerio, Mara et surtout Simone, dont le bureau est devenu depuis l'an dernier l'endroit le plus accueillant de toute la Fac d'Orsay, et merci également à Andrea, Federico et Mattia, qu'en ces derniers temps j'ai fréquenté beaucoup trop peu.
Et j'en suis finalement à mes plus chers amis, sans lesquels la vie serait beaucoup moins intéressante. Que ferais-je sans quelqu'un qui me prévienne à la une du matin qu'il ne faudra peut-être pas attendre l'année 3072 pour le film de WoW (cela te dit quelque chose, peut-être, Bonvi) ? Sans Danieluccio et son examen de stats à valider, sans Manuel-qui-cet-été-ne-travaille-absolument-pas ? Sans Monsieur le Président et sa Première Dame, sans Jessica, Simone, Carmen, Monica, Desi, Matte, Ste, Fra et Adri, ces années auraient été bien différentes et sans doute bien moins colorées et joyeuses.
Je suis énormément reconnaissant envers mes parents Marina et Sabino (mamma e papà !) pour tout ce qu'ils ont fait pour moi et pour avoir soutenu mes choix, même quand ils auraient sans doute préféré que je choisisse un autre chemin. Per ringraziare Nonna Jole tornerò invece alla mia lingua madre (prima di sentirmi dire che "c'erano parole strane") : il suo contributo a questa tesi non va affatto trascurato, dal momento che è stata lei, tanti anni fa, ad insegnarmi a contare i punti, giocando a carte insieme. Per questo, e per tutto quello che ha fatto per me, le dico GRAZIE.
Et enfin, ma maîtrise du français n'est sans doute pas suffisante pour exprimer mon immense gratitude à Alessandra, qui a toujours été là pour moi. Tout simplement merci, merci et encore mille fois merci.

Introduction

Soient K un corps de nombres et A/K une variété abélienne dont nous notons g la dimension. Pour tout nombre premier ℓ et tout entier positif n, notons $A[\ell^n] \subseteq A(\overline{K})$ le groupe des points de torsion de A dont l'ordre divise ℓ^n. Comme A est définie sur K, l'action naturelle du groupe de Galois absolu $\mathrm{Gal}(\overline{K}/K)$ sur $A(\overline{K})$ stabilise $A[\ell^n]$ en tant que ensemble, ce qui nous fournit une représentation continue

$$\rho_{\ell^n} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\left(A[\ell^n]\right).$$

C'est à ces représentations galoisiennes, et à leur variantes ℓ-adiques que nous rappellerons dans un instant, que l'on s'intéresse dans toute la suite. Remarquons que le corps $K(A[\ell^n])$ obtenu en adjoignant à K les coordonnées des points de $A[\ell^n]$ est une extension galoisienne de K : il est alors clair sur la définition que l'image G_{ℓ^n} de ρ_{ℓ^n} s'identifie au groupe de Galois de $K(A[\ell^n])$ sur K. De plus, il est bien connu que – K étant de caractéristique 0 – le groupe $A[\ell^n]$ est un $\mathbb{Z}/\ell^n\mathbb{Z}$-module libre de rang $2g$.
Pour ℓ fixé, les groupes finis $A[\ell^n]$ forment un système projectif pour lequel les morphismes de transition sont donnés par la multiplication par ℓ : la limite inverse de ce système est appelée le module de Tate ℓ-adique de A, souvent noté $T_\ell A$. Comme chaque cran fini $A[\ell^n]$ est libre de rang $2g$ sur $\mathbb{Z}/\ell^n\mathbb{Z}$, on voit aisément que $T_\ell A$ est un $\mathbb{Z}_\ell$-module libre de rang $2g$, et nous serons très souvent amenés à fixer une $\mathbb{Z}_\ell$-base de $T_\ell A$, ce qui nous permettra d'écrire certaines égalités "en coordonnées". Remarquons dès à présent que la construction du module de Tate s'applique également à tout groupe ℓ-divisible, et on peut notamment considérer le module de Tate du groupe multiplicatif $\mathbb{G}_m$, un objet qui sera utile par la suite.

Il n'est pas difficile de vérifier que les représentations (ρ_{ℓ^n}), quand n varie, forment à leur tour un système projectif, dont la limite ρ_{ℓ^∞} est donc une application continue

$$\rho_{\ell^\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(T_\ell A)$$

qui est traditionnellement appelée la *représentation ℓ-adique associée à A* et dont nous désignons l'image par G_{ℓ^∞}.

Les représentations ρ_{ℓ^∞} ainsi construites jouent un rôle très important dans la théorie des nombres contemporaine : il suffira par exemple de rappeler que la première preuve de la conjecture de Mordell, donnée par Faltings dans les années '80 [26], repose très fortement sur leur étude, et qu'elles interviennent dans la preuve du dernier théorème de Fermat [142] [132].

Rappelons dès maintenant une des propriétés fondamentales des représentations ρ_{ℓ^∞}, dont on fera un usage extensif par la suite : si on désigne par $A^\vee$ la duale de A, il existe sur $T_\ell(A) \times T_\ell(A^\vee)$ une

forme bilinéaire non-dégénérée $\langle \cdot, \cdot \rangle$, dite *accouplement de Weil*, prenant ses valeurs dans le groupe $\varprojlim_n \mu_{\ell^n}(\overline{K}) = T_\ell \mathbb{G}_m$ et Galois-équivariante, au sens où on a l'égalité

$$\langle \rho_{\ell^\infty}(g)v, \rho_{\ell^\infty}(g)w \rangle = \chi_\ell(g) \cdot \langle v, w \rangle$$

pour tout $v \in T_\ell(A), w \in T_\ell(A^\vee)$ et tout $g \in \mathrm{Gal}(\overline{K}/K)$.
Quand la variété A est présentée comme une jacobienne (notamment, si elle est une courbe elliptique) on dispose de plus d'une polarisation principale canonique, à savoir un K-isomorphisme f entre A et $A^\vee$, et l'accouplement de Weil peut alors se réinterpréter comme une forme non-dégénérée, bilinéaire, antisymétrique et Galois-équivariante sur $T_\ell A$ que notons encore $\langle \cdot, \cdot \rangle$:

$$T_\ell(A) \times T_\ell(A) \xrightarrow{\mathrm{id} \times f} T_\ell A \times T_\ell A^\vee \xrightarrow{\langle \cdot, \cdot \rangle} T_\ell \mathbb{G}_m. \tag{1}$$

L'existence d'une telle forme antisymétrique fait que l'image de ρ_{ℓ^∞} est contenue dans le groupe des similitudes symplectiques de la forme $\langle \cdot, \cdot \rangle$, groupe que nous notons $\mathrm{GSp}\,(T_\ell A, \langle \cdot, \cdot \rangle)$, ou même simplement $\mathrm{GSp}\,(T_\ell A)$, voire encore $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ si le choix d'une $\mathbb{Z}_\ell$-base de $T_\ell A$ a été fait.
Dans le cas général, le choix d'une K-polarisation quelconque $f : A \to A^\vee$ induit une forme $T_\ell(A) \times T_\ell(A) \to T_\ell \mathbb{G}_m$ comme en (1) : pour tout premier ℓ ne divisant pas le degré de f on a $T_\ell A \cong T_\ell A^\vee$, d'où encore une inclusion $G_{\ell^\infty} \subseteq \mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$, alors que pour les premiers divisant $\deg f$ on a seulement $G_{\ell^\infty} \subseteq \mathrm{GSp}_{2g}(\mathbb{Q}_\ell)$.
Finalement, sur toutes ces questions on peut aussi prendre un point de vue adélique, ce qui nous amène très naturellement à introduire la *représentation adélique associée à* A, à savoir le produit des ρ_{ℓ^∞} pour tous les premiers ℓ :

$$\rho_\infty : \mathrm{Gal}(\overline{K}/K) \xrightarrow{\prod \rho_{\ell^\infty}} \prod_\ell \mathrm{Aut}(T_\ell A). \tag{2}$$

Pour mieux préciser le cadre dans lequel se situe ce travail il est aussi indispensable de rappeler la conjecture de Mumford-Tate (cf. la conjecture 1 ci-après). Pour comprendre comment cette conjecture apparaît de manière naturelle à partir de la philosophie motivique de Grothendieck, remarquons d'abord que, pour chaque premier ℓ, le module de Tate $T_\ell A$ peut s'identifier au dual de $H^1_{\acute{e}t}\left(A_{\overline{K}}, \mathbb{Z}_\ell\right)$, ce qui nous permet d'interpréter les représentations ρ_{ℓ^∞} comme étant données par l'action naturelle de $\mathrm{Gal}(\overline{K}/K)$ sur le H^1 étale de A. Or la philosophie motivique prédit que la cohomologie étale n'est que une incarnation d'une cohomologie universelle, qui admet aussi une réalisation (de Betti) comme cohomologie habituelle de l'espace topologique $A(\mathbb{C})$. On va donc examiner quelques propriétés de cette cohomologie, qui dans le cas d'une variété abélienne possède une description particulièrement simple : comme $A(\mathbb{C})$ est topologiquement un tore, la formule de Künneth implique que l'algèbre de cohomologie $H^\bullet\left(A(\mathbb{C}), \mathbb{Z}\right)$ est canoniquement isomorphe à l'algèbre extérieure sur $H^1(A(\mathbb{C}), \mathbb{Z}) \cong H_1(A(\mathbb{C}), \mathbb{Z})^\vee$, et d'après la théorie classique (qui remonte essentiellement à Riemann) le sous-réseau $H_1(A(\mathbb{C}), \mathbb{Z})$ de l'espace vectoriel complexe $H_1(A(\mathbb{C}), \mathbb{C})$ décrit complètement la variété abélienne complexe $A(\mathbb{C})$. D'autre part, comme toute variété abélienne est en particulier projective, et donc Kählerienne, la théorie de Hodge nous donne des renseignements plus précis sur cette cohomologie : on dispose en effet d'une décomposition canonique

$$H^n\left(A(\mathbb{C}), \mathbb{C}\right) \cong \bigoplus_{p+q=n} H^{p,q},$$

où $H^{p,q}$ est donné par $H^{p,q} = H^p\left(A(\mathbb{C}), \Omega^q\right)$ et respecte $\overline{H^{p,q}} = H^{q,p}$. Par définition, cela signifie que le $\mathbb{Q}$-espace vectoriel $H^n(A(\mathbb{C}), \mathbb{Q})$ est une *structure de Hodge* (pure de poids n). D'après la description des structures de Hodge donnée par Deligne [23], on sait que une structure de Hodge sur le $\mathbb{Q}$-espace vectoriel V est équivalent à la donnée d'un morphisme de groupes algébriques

$$h : \mathbb{S} \to \mathrm{GL}_{V_{\mathbb{R}}},$$

où $\mathbb{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\left(\mathbb{G}_{m,\mathbb{C}}\right)$ est le "tore de Deligne", et $V_{\mathbb{R}}$ désigne le $\mathbb{R}$-espace vectoriel $V \otimes_{\mathbb{Q}} \mathbb{R}$. La $\mathbb{Q}$-fermeture de Zariski de l'image de h est alors appelée le *groupe de Mumford-Tate* de la structure de Hodge V. Dans le cas des variétés abéliennes, la formule de Künneth susmentionnée fait que la décomposition de Hodge de $H^n(A(\mathbb{C}), \mathbb{C})$ s'obtient à partir de celle de $H^1(A(\mathbb{C}), \mathbb{C})$, et il est donc spécialement intéressant d'étudier cette dernière structure de Hodge (de poids 1).
Soit donc A une variété abélienne définie sur corps K, de caractéristique 0 et finiment engendré sur son corps premier. Pour tout plongement σ de K dans $\mathbb{C}$ on obtient une variété abélienne complexe $A_\sigma := A \times_\sigma \mathbb{C}$, et on définit le groupe de Mumford-Tate de A (noté $\mathrm{MT}(A)$) comme le groupe de Mumford-Tate de la structure de Hodge $H^1\left(A_\sigma(\mathbb{C}), \mathbb{Q}\right)$. Grâce au fait que tous les cycles de Hodge sont absolument de Hodge ([23, Theorem 2.11]), cette définition ne dépend pas de σ, de sorte que si A est définie sur un corps K de type fini sur $\mathbb{Q}$ (notamment un corps de nombres) on peut parler du groupe de Mumford-Tate de A sans spécifier un plongement de K dans $\mathbb{C}$. Remarquons aussi que $\mathrm{MT}(A)$ est un objet purement géométrique, et en particulier invariant par extension du corps de base K ; de plus, il ne dépend que de la $\mathbb{C}$-classe d'isogénie de A, car en effet la même propriété est vraie pour $H^1(A(\mathbb{C}), \mathbb{Q})$. Finalement, il n'est pas difficile de voir que le groupe $\mathrm{MT}(A)$ est contenu dans $\mathrm{GSp}_{2g,\mathbb{Q}}$: en effet, toute polarisation $\varphi : A \to A^\vee$ induit une forme bilinéaire alternée

$$H^1(A(\mathbb{C}), \mathbb{Q}) \times H^1(A(\mathbb{C}), \mathbb{Q}) \xrightarrow{\mathrm{id} \times \varphi} H^1(A(\mathbb{C}), \mathbb{Q}) \times H^1(A^\vee(\mathbb{C}), \mathbb{Q}) \to \mathbb{Q}$$

qui est une polarisation au sens des structures de Hodge, et il est bien connu que cela entraîne l'inclusion $\mathrm{MT}(A) \subseteq \mathrm{GSp}_{2g,\mathbb{Q}}$. Par ailleurs, remarquons que la dualité de Poincaré nous permet d'identifier $H_1(A(\mathbb{C}), \mathbb{Q})$ avec $H^1(A(\mathbb{C}), \mathbb{Q})^\vee$, ce qui induit une $\mathbb{Q}$-structure de Hodge sur $H_1(A(\mathbb{C}), \mathbb{Q})$ dont le groupe de Mumford-Tate coïncide avec celui de $H^1(A(\mathbb{C}), \mathbb{Q})$.

Revenons maintenant à nos représentations galoisiennes ; on suppose à nouveau que A est définie sur un corps de nombres K. D'après le théorème de comparaison en cohomologie étale, on dispose d'un isomorphisme

$$H^1(A(\mathbb{C}), \mathbb{Q}) \otimes \mathbb{Q}_\ell \cong H^1_{\acute{e}t}\left(A_{\overline{K}}, \mathbb{Q}_\ell\right)$$

qui s'étend d'ailleurs à toute l'algèbre de cohomologie, à savoir on dispose plus généralement d'isomorphismes

$$H^\bullet_{\acute{e}t}\left(A_{\overline{K}}, \mathbb{Q}_\ell\right) \cong \Lambda^\bullet H^1_{\acute{e}t}\left(A_{\overline{K}}, \mathbb{Q}_\ell\right) \cong \Lambda^\bullet H^1\left(A(\mathbb{C}), \mathbb{Q}\right) \otimes \mathbb{Q}_\ell.$$

On est alors amené à comparer $\mathrm{MT}(A)$ – défini en termes de la réalisation Betti de A – avec un objet correspondant sur le côté galoisien : c'est dans cet esprit qu'on définit le *groupe de monodromie ℓ-adique* de A, souvent noté $\mathcal{G}_\ell(A)$, comme la $\mathbb{Q}_\ell$-clôture de Zariski dans $\mathrm{GL}_{T_\ell(A)\otimes\mathbb{Q}_\ell}$ de l'image de la représentation $\rho_{\ell\infty}$ introduite en (2). Le lecteur averti pourrait maintenant remarquer une légère asymétrie dans nos définitions, car $T_\ell(A) \otimes \mathbb{Q}_\ell$ s'identifie plutôt au *dual* de $H^1(A(\mathbb{C}), \mathbb{Q}) \otimes \mathbb{Q}_\ell$: toutefois, cet asymétrie disparaît lorsque on remarque que – comme on l'a déjà rappelé – le groupe

de Mumford-Tate de $H^1(A(\mathbb{C}),\mathbb{Q})$ et celui de $H_1(A(\mathbb{C}),\mathbb{Q})$ peuvent être identifiés. Finalement, l'isomorphisme canonique $H_1(A(\mathbb{C}),\mathbb{Q})\otimes\mathbb{Q}_\ell \cong T_\ell(A)\otimes\mathbb{Q}_\ell$ nous permet, par extension des scalaires à $\mathbb{Q}_\ell$, de considérer $\mathrm{MT}(A)\times_{\mathbb{Q}}\mathbb{Q}_\ell$ comme un sous-groupe de $\mathrm{GL}_{T_\ell(A)\otimes\mathbb{Q}_\ell}$. Il est alors très tentant de conjecturer que l'on devrait avoir $\mathcal{G}_\ell(A)\cong\mathrm{MT}(A)\times_{\mathbb{Q}}\mathbb{Q}_\ell$ en tant que sous-groupes de $\mathrm{GL}_{T_\ell(A)\otimes\mathbb{Q}_\ell}$, mais cette conjecture naïve est fausse : en effet, le groupe $\mathcal{G}_\ell(A)$ n'est pas connexe en général, alors que le groupe $\mathrm{MT}(A)$ est connexe par définition. Toutefois, un célèbre théorème de Serre ([121, § 2.2.3]) nous assure que – quitte à remplacer K par une extension finie – on peut garantir que $\mathcal{G}_\ell(A)$ est connexe pour tout ℓ ; de plus, il est clair que remplacer K par une extension finie ne change pas la composante neutre $\mathcal{G}_\ell(A)^0$ (car tout sous-groupe fermé de G_{ℓ^∞} d'indice fini est automatiquement dense dans $\mathcal{G}_\ell(A)^0$ pour la topologie de Zariski). Ce cercle d'idées apporta Mumford et Tate [84] à formuler la célèbre conjecture :

Conjecture 1. *(Mumford-Tate) Pour tout corps de nombres K et toute variété abélienne A/K on a l'égalité $\mathcal{G}_\ell(A)^0=\mathrm{MT}(A)\times_{\mathbb{Q}}\mathbb{Q}_\ell$.*

Beaucoup de progrès ont été faits en direction de cette conjecture, mais dans le cas général elle reste encore largement ouverte. Le résultat le plus général à ce propos à été prouvé par Borovoĭ [14], Deligne [23, Exp. I, 2.9, 2.11], et Pjateckiĭ-Šapiro [100], qui montrent

Théorème 2. *Pour toute variété abélienne A sur un corps de nombres et tout premier ℓ on a $\mathcal{G}_\ell(A)^0\subseteq\mathrm{MT}(A)\times_{\mathbb{Q}}\mathbb{Q}_\ell$.*

A la lumière de ce théorème on voit que la partie encore ouverte de la conjecture de Mumford-Tate revient essentiellement à dire que l'image G_{ℓ^∞} de la représentation ρ_{ℓ^∞} est "aussi grosse que possible", à savoir suffisamment grosse pour que sa fermeture de Zariski coïncide avec $\mathrm{MT}(A)\times_{\mathbb{Q}}\mathbb{Q}_\ell$: pour cette raison on désigne souvent les résultats dans cette veine par le nom de "théorèmes de l'image ouverte".

D'après les travaux de plusieurs mathématiciens (notamment Serre, Pink, Ribet, Chi, Tanke'ev, Banaszak, Gajda, Krasoń, Hall...) on connaît des tels résultats pour des nombreuses classes de variétés abéliennes : sans prétendre à l'exhaustivité, rappelons ici que la conjecture de Mumford-Tate a été prouvé pour les variétés A de dimension impaire satisfaisant à $\mathrm{End}_{\overline{K}}(A)=\mathbb{Z}$ (Serre [118]), pour les variétés satisfaisant à $\mathrm{End}_{\overline{K}}(A)=\mathbb{Z}$ et dont la dimension est en dehors d'un ensemble de densité zéro (Pink [98]), pour les variétés de type CM (Shimura-Taniyama [126], Pohlmann [101], Serre-Tate [124]) et pour les variétés dénommées "de type GL_2" (Ribet [109]).

On peut aussi donner des résultats qui font intervenir un invariant plus fin que la simple dimension, à savoir la *dimension relative*. Rappelons que si A est une variété abélienne géométriquement simple, alors l'algèbre $\mathrm{End}_{\overline{K}}(A)\otimes\mathbb{Q}$ est un corps gauche D (de plus, il s'agit d'une algèbre admettant une involution positive), dont le centre est un corps de nombres E. On distingue alors le type de A selon le type de l'algèbre D dans la classification donnée par Albert [1] [2], et on définit la dimension relative (qui est toujours un nombre entier) par la formule

$$\dim\text{ rél}(A)=\begin{cases}\frac{\dim A}{[E:\mathbb{Q}]\sqrt{[D:E]}}, & \text{si } A \text{ est de type I, II ou III}\\ \frac{2\dim A}{[E:\mathbb{Q}]\sqrt{[D:E]}}, & \text{si } A \text{ est de type IV.}\end{cases}$$

A titre d'exemple on a alors le résultat suivant :

Théorème 3. *(Banaszak, Gajda, Krasoń [6]) La conjecture de Mumford-Tate est vraie pour toute variété abélienne simple de type I ou II et de dimension relative impaire.*

Bien que ces résultats soient très satisfaisants du point de vue de la conjecture de Mumford-Tate, ils ne répondent pas à la question de *décrire* précisément les images G_{ℓ^∞} des représentations ρ_{ℓ^∞} : la propriété d'être ouvert est invariante par passage à un sous-groupe d'indice fini, donc ces théorèmes ne fournissent pas un renseignement très précis sur G_{ℓ^∞} lui-même.
Dans certains cas particuliers on dispose toutefois d'une description de G_{ℓ^∞}, du moins pour ℓ suffisamment grand. Par exemple, dans son cours au Collège de France de 1986 Serre a prouvé le théorème suivant :

Théorème 4. *Soit A/K une variété abélienne de dimension 2, 6, ou un nombre impair. Supposons que $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$: alors l'image G_∞ de la représentation adélique ρ_∞ est ouverte dans le produit restreint $\prod'_\ell \mathrm{GSp}_{2\dim A}(\mathbb{Q}_\ell)$. En particulier, il existe un entier $\ell_0(A,K)$ tel que pour tout $\ell > \ell_0(A,K)$ on a $G_{\ell^\infty} = \mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$.*

Ce théorème, qui généralise un résultat précédent concernant les courbes elliptiques ([116]), n'est malheureusement pas effectif : la preuve ne fournit aucune indication sur la valeur de $\ell_0(A/K)$.

L'extension engendrée par un point de torsion

Les différents théorèmes de l'image ouverte permettent aussi d'étudier la question suivante :

Problème 5. *Soit A/K une variété abélienne et $P \in A_{tors}(\overline{K})$ un point de torsion d'ordre N. Quel est la relation entre le degré $[K(P) : K]$ et l'ordre N ?*

Ce genre de question apparaît naturellement par exemple dans certains problèmes de type Manin-Mumford relatif (cf. [74], pour ne donner qu'un exemple), et est bien sûr très intéressant en soi, surtout en relation avec la conjecture nommée "de la torsion uniforme" :

Conjecture 6. *Soient K un corps de nombres et g un entier positif. Existe-t-elle une constante $C(K,g)$ telle que pour toute variété abélienne A définie sur K et de dimension g on a*

$$|A_{tors}(K)| \leq C(K,g)?$$

Plusieurs mathématiciens se sont occupés de différentes variantes du problème 5 ; notamment, en utilisant la théorie de la transcendance, Bertrand [9] a prouvé

Théorème 7. *Soit A/K une variété abélienne de dimension g sur un corps de nombres K. Alors pour tout $\varepsilon > 0$ il existe une constante $C(A,K,\varepsilon)$ telle que, pour tout point $P \in A_{tors}(\overline{K})$ d'ordre N, on a $[K(P) : K] \geq C(A,K,\varepsilon)N^{1/(2+g+\varepsilon)}$. La constante $C(A,K,\varepsilon)$ est ici effective en fonction de ε et des équations définissant A.*

D'après le travaux de Serre on sait toutefois qu'un tel résultat est relativement loin d'être optimal, comme il le montre le théorème suivant :

Théorème 8. *([118]) Soit A/K une variété abélienne sur un corps de nombres K. Supposons que aucun facteur géométriquement simple de $A_{\overline{K}}$ n'est de type CM : alors pour tout $\varepsilon > 0$ il existe une constante $C(A, K, \varepsilon)$ telle que, pour tout point $P \in A_{tors}(\overline{K})$ d'ordre N, on a*

$$[K(P) : K] \geq C(A, K, \varepsilon) N^{2-\varepsilon}.$$

Si $A_{\overline{K}}$ a un facteur avec multiplication complexe, alors le même énoncé reste vrai quitte à remplacer $2 - \varepsilon$ par $1 - \varepsilon$.

L'approche de Serre repose sur l'étude des représentations ρ_{ℓ^∞}, mais est une fois de plus ineffectif, au sens où étant donnés des équations pour A la preuve du théorème 8 ne fournit aucune indication sur la valeur de $C(A, K, \varepsilon)$. Une exception est constituée par les variétés avec multiplication complexe, pour lesquelles Silverberg [128] – en s'appuyant sur les travaux fondamentaux de Shimura-Taniyama [126] et Serre-Tate [124] – a montré le résultat suivant, qui a de plus l'avantage d'être *uniforme* en A :

Théorème 9. *Soient K un corps de nombres et g un entier positif. Pour tout $\varepsilon > 0$ il existe une constante effective $C(K, g, \varepsilon)$ avec la propriété suivante : pour toute variété abélienne A/K de dimension g avec multiplication complexe, et pour tout point $P \in A_{tors}(\overline{K})$ d'ordre N, on a $[K(P) : K] \geq C(K, g, \varepsilon) N^{1-\varepsilon}$.*

Tous ces problèmes admettent évidemment des variantes, également intéressantes, qui concernent par exemple l'extension engendrée par *tous* les points de N-torsion d'une variété A. Une petite réflexion montre que ces variantes reviennent essentiellement à une reformulation dans ces termes de résultats du type "image ouverte" ; on mentionne à ce propos un théorème de Ribet [110], qui donne un tel résultat pour les variétés CM :

Théorème 10. *Soit A/K une variété abélienne de type CM. Il existe des constantes positives C_1, C_2, dépendantes de A et de K, et un entier $r > 0$, également dépendant de A, tels que pour tout N entier positif on a*

$$C_1 N^r \omega(N)^r \leq [K(A[N]) : K] \leq C_2 N^r \omega(N)^r,$$

où $\omega(N)$ est le nombre de facteurs premiers distincts de N.

Remarquons que aussi ce résultat est non-effectif en ce qui concerne les constantes C_1 et C_2.

Finalement, le même genre de techniques permet aussi d'étudier une question complémentaire à celles indiquées ci-dessus : si A/K est encore une fois une variété abélienne fixée, on peut essayer de comprendre comment varie le groupe de torsion $A_{\text{tors}}(K')$ lorsque l'on fait varier K' parmi les extensions finies de K. C'est le but des récents travaux de Hindry et Ratazzi [38] [39] [40], qui introduisent l'invariant suivant :

Définition 11. Soit K un corps de nombres et A/K une variété abélienne. On pose

$$\gamma(A) = \inf \left\{ x > 0 \;\middle|\; \exists C > 0 \quad \forall K'/K \text{ finie}, \quad \left|A_{\text{tors}}(K')\right| \leq C[K' : K]^x \right\}.$$

En s'appuyant sur des résultats du type "image ouverte", Hindry et Ratazzi calculent l'invariant $\gamma(A)$ pour des nombreuses classes des variété abéliennes ; toute forme effective des théorèmes de

l'image ouverte permettrait en particulier de mieux comprendre la nature de la constante C, et plus généralement toute nouvelle instance (même non effective) de la conjecture de Mumford-Tate permettrait sans doute d'élargir la classe des variétés abéliennes pour lesquelles on sait déterminer cet exposant optimal $\gamma(A)$.

Questions d'effectivité

Les résultats sur les représentations galoisiennes rappelés aux paragraphes précédents montrent très clairement une tendance assez générale : même si on a beaucoup de renseignements qualitatifs sur les images G_{ℓ^∞} des représentations ρ_{ℓ^∞}, presque aucun d'entre eux n'est effectif. Dans cette thèse on se propose notamment de donner des versions effectives de certains des énoncés mentionnés aux paragraphes précédents : concrètement, il s'agit de donner des valeurs pour les différentes bornes qui aient une dépendance aussi simple que possible en les données A et K. D'après les travaux de Faltings [26], on possède un moyen canonique de mesurer la complexité arithmétique d'une variété abélienne : sa hauteur (de Faltings) stable $h(A)$. On sait de plus que – pour tout corps K fixé et tout $B \in \mathbb{R}$ – il n'y a que un nombre fini de variétés abéliennes A/K de hauteur bornée par B : il est donc évident que les différentes bornes apparaissant dans les différents théorèmes de l'image ouverte doivent pouvoir s'exprimer en fonction de $h(A)$. On doit également choisir une mesure de complexité pour les corps de nombres K : traditionnellement, on utilise à cet effet le discriminant $\Delta_{K/\mathbb{Q}}$, mais on verra que très souvent le degré $[K : \mathbb{Q}]$ sera un invariant suffisant pour caractériser la dépendance en K. Nous tenons à souligner dès maintenant que nos résultats faisant apparaître la hauteur de Faltings de A ont une dépendance en $h(A)$ qui est *polynomiale*, et qu'il en va de même pour la dépendance en $[K : \mathbb{Q}]$: ceci est rendu possible par l'utilisation d'une forme explicite très précise du théorème dit "d'isogénie", originairement prouvé dans une forme non effective par Faltings [26] et ensuite redémontré de façon effective par Masser et Wüstholz grâce à des techniques de théorie de la transcendance [70] [72] [73]. La version de ce théorème que nous utilisons ici est plutôt due à Gaudron et Rémond [28] ; pour énoncer leur résultat, nous introduisons la définition suivante :

Définition 12. On pose

$$b(d, g, h) := \left((14g)^{64g^2} d \max\{1, h, \log d\}\right)^{2^{10} g^3}.$$

Plusieurs de nos résultats s'exprimeront à l'aide de la fonction $b(d, g, h)$. Avec cette notation, l'un des théorèmes principaux de [28] s'énonce ainsi :

Théorème 13. *Soit A/K une variété abélienne de dimension g. Pour toute variété abélienne A^* définie sur K qui est K-isogène à A, il existe une K-isogénie $A^* \to A$ de degré borné par $b([K : \mathbb{Q}], g, h(A))$.*

Contenu de la thèse

Les résultats rappelés dans les paragraphes précédents nous amènent à considérer des questions que l'on peut grossièrement distinguer en trois catégories :

1. image ouverte : montrer que $\mathcal{G}_\ell(A)^0 = \mathrm{MT}(A) \times_{\mathbb{Q}} \mathbb{Q}_\ell$ pour des nouvelles classes de variétés abéliennes.

2. effectivité pour ℓ suffisamment grand : étant donné une variété abélienne A/K, donner une borne effective $\ell_0 = \ell_0(A/K)$ telle que, pour tout $\ell > \ell_0$, on sait décrire explicitement le groupe G_{ℓ^∞}.

3. effectivité ℓ-adique et adélique : étant donné A/K et un premier ℓ, calculer une borne $c(A/K, \ell)$ sur les indices $[\mathcal{G}_\ell(A)(\mathbb{Z}_\ell) : G_{\ell^\infty}]$ et $[\prod_\ell \mathcal{G}_\ell(A)(\mathbb{Z}_\ell) : G_\infty]$, si ce dernier est fini.

Dans cette thèse nous présentons quelques résultats originaux en direction des questions 1, 2 et 3 ci-dessus. Les différents chapitres sont rédigés sous la forme d'articles essentiellement indépendants les uns des autres ; ainsi, le lecteur remarquera que certaines parties des introductions sont répétées.

La première (et plus longue) partie de la thèse est consacrée au problème de rendre complètement effectifs les théorèmes du type "image ouverte" pour les variétés abéliennes, explicitant aussi toutes les constantes qui interviennent.

Nous traitons d'abord (dans le chapitre 1) le cas des courbes elliptiques n'ayant pas de multiplication complexe. Précisément, en passant par un résultat de structure pour certains algèbres de Lie entières (à coefficients dans $\mathbb{Z}_\ell$), nous montrons notamment des versions effectives des théorèmes 4 et 8 pour $g = 1$:

Théorème 14. *Soit E/K une courbe elliptique sans CM. On a*

$$\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right] < C_1 \cdot [K:\mathbb{Q}]^{C_2} \cdot \max\left\{1, h(E), \log[K:\mathbb{Q}]\right\}^{2C_2},$$

où $C_1 = \exp(6 \cdot 10^{29527})$ et $C_2 = 4.9 \cdot 10^{10}$.
Si $P \in E_{\mathrm{tors}}(\overline{K})$ est un point de torsion d'ordre N on a

$$[K(P):K] \geq \left(\zeta(2) \cdot \left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right]\right)^{-1} \cdot N^2.$$

La preuve de ces résultats répose sur le théorème d'isogénie, combiné avec une extension de la théorie des algèbres de Lie entières dûe à Pink [97]. Rappelons brièvement cette construction. Soit A un anneau semi-local, et soit I l'intersection de tous les idéaux maximaux de A. On suppose que l'anneau A est complet pour la topologie I-adique, et que A/I est annulé par un premier p, supposé *impair*. La théorie de Pink montre alors qu'il a une corréspondence bijective et fonctorielle entre les sous-groupes pro-p de $\mathrm{SL}_2(A)$ et les couples (L, Δ), où L est une sous-algèbre de Lie de $\mathfrak{sl}_{2,A}$ qui vérifie $\cap_{n\geq 0} L^n = \{0\}$ et $\mathrm{tr}(L \cdot L) \cdot L \subseteq L$, et Δ est un sous-ensemble fermé de $L/[L,L]$ qui respecte certaines propriétés additionnelles. De plus, si (L, Δ) est le couple qui corréspond au groupe G, alors la connaissance de L est suffisante à déterminer le sous-groupe dérivé de G.
On aimerait pouvoir étudier l'image de Galois par l'intermédiaire de cette correspondence : en effet, caractériser les sous-algèbres de Lie de $\mathfrak{sl}_{2,\mathbb{Z}_\ell}$ est essentiellement un problème d'algèbre linéaire, donc rélativament facile, alors que travailler directement avec les sous-groupes de $\mathrm{SL}_2(\mathbb{Z}_\ell)$ est (à priori) bien plus compliqué. Malheuresement, la théorie de Pink n'est pas suffisante pour nos applications, pour trois raisons différentes :

1. elle ne s'applique qu'aux sous-groupes de $\mathrm{SL}_2(\mathbb{Z}_\ell)$, alors que l'image de Galois est un sous-groupe de $\mathrm{GL}_2(\mathbb{Z}_\ell)$;

2. elle ne s'étend pas au cas $p = 2$;

3. son domaine d'applicabilité est limité aux sous-groupes de $\mathrm{SL}_2(\mathbb{Z}_\ell)$ qui sont de plus pro-ℓ, ce qui en général n'est pas le cas pour nos groupes G_{ℓ^∞}.

Le premier problème n'est pas très difficile à contourner en considérant le groupe dérivé de G_{ℓ^∞} (qui est automatiquement un sous-groupe de $\mathrm{SL}_2(\mathbb{Z}_\ell)$), mais les autres deux obstacles sont plus sérieux. Il y a également une solution rélativement simple (mais pas complètement satisfaisante) pour le troisième problème : si on remplace le corps K par une extension convenable, on peut toujours supposer que – pour un premier ℓ fixé – le groupe G_{ℓ^∞} est un groupe pro-ℓ. Malheuresement, le degré de cette extension varie en général avec ℓ, et cela se traduirait par un résultat final qui n'est pas polynomiale en la hauteur de E.
Nous résolvons complètement les problèmes 2 et 3 par les résultats suivants, qui peuvent être considérés comme une éxtension de la théorie de Pink dans le cas où l'anneau A est $\mathbb{Z}_p$ (où p est n'importe quel nombre premier, y compris $p = 2$).

Théorème 15. *Soit ℓ un nombre premier impair (resp. $\ell = 2$). Pour tout sous-groupe fermé G de $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (resp. pour tout sous-groupe fermé dont la réduction modulo 2 est triviale si $\ell = 2$), soit $L(G)$ le $\mathbb{Z}_\ell$-sous-module de $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ engendré par l'ensemble $\left\{ g - \frac{\mathrm{tr}(g)}{2} \cdot \mathrm{Id} \;\middle|\; g \in G \right\}$.*
Soit H un sous-groupe fermé de $\mathrm{GL}_2(\mathbb{Z}_\ell)$. Il existe un sous-groupe H_1 de H, d'indice au plus 24 (resp. un sous-groupe d'indice au plus 192 ayant de plus une réduction triviale modulo 2, quand $\ell = 2$), tel que l'on a l'implication suivante pour tout entier positif s : si $L(H_1)$ contient $\ell^s \mathfrak{sl}_2(\mathbb{Z}_\ell)$, alors H_1 contient

$$\mathcal{B}_\ell(4s) := \left\{ g \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; g \equiv \mathrm{Id} \pmod{\ell^{4s}} \right\} \quad \textit{(resp. } \mathcal{B}_2(6s) \textit{ pour } \ell = 2\textit{)}.$$

Remarquons que le résultat pour $\ell = 2$ est complètement indépendant des résultats de Pink. Signalons également que nous construisons des exemples qui montrent qu'on ne peut pas éviter de remplacer H par un sous-groupe, de sorte que la forme de ce résultat est essentiellement optimale.

Dans le chapitre 2 nous étendons ensuite la méthode et les résultats du chapitre précédent au cas d'un produit arbitraire de courbes elliptiques sans multiplication complexe :

Théorème 16. *Soit $n \geq 2$ et soient $E_1, \ldots, E_n$ des courbes elliptiques définies sur un corps de nombres K, deux à deux non isogènes sur $\overline{K}$. Supposons que $\mathrm{End}_{\overline{K}}(E_i) = \mathbb{Z}$ pour $i = 1, \ldots, n$, et notons G_∞ l'image de $\mathrm{Gal}(\overline{K}/K)$ dans*

$$\prod_\ell \mathrm{Aut}(T_\ell(E_1)) \times \cdots \times \mathrm{Aut}(T_\ell(E_n)) \subset \mathrm{GL}_2(\hat{\mathbb{Z}})^n.$$

Soit $\gamma := 10^{13}, \delta := \exp\exp\exp(13)$, et $H = \max\{1, \log[K : \mathbb{Q}], \max_i h(E_i)\}$. Le groupe G_∞ a indice au plus

$$\delta^{n(n-1)} \cdot \left([K : \mathbb{Q}] \cdot H^2 \right)^{\gamma n(n-1)}$$

dans

$$\Delta := \left\{ (x_1, \ldots, x_n) \in \mathrm{GL}_2(\hat{\mathbb{Z}})^n \;\middle|\; \det x_i = \det x_j \;\; \forall i, j \right\}.$$

La preuve de ce théorème nécessite d'une étude assez fine de certains sous-groupes fermés de $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$. Il s'agit d'une exténsion ultérieure des résultats de Pink au cas où l'anneau de base est $\mathbb{Z}_\ell^n$; même si le résultat pour $n = 2$ nous suffirait, nous traitons le cas d'un n arbitraire dans le chapitre 8.

Dans le chapitre 3 nous commençons à considérer certains cas de dimension supérieure : nous y étudions notamment les représentations ρ_{ℓ^∞} associées à des surfaces abéliennes A géométriquement simples. Rappelons que – selon l'arithmétique de l'algèbre des endomorphismes de A – on distingue 4 types de surfaces : génériques ($\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$), à multiplication réelle ($\mathrm{End}_{\overline{K}}(A)$ est un ordre dans un corps quadratique réel), à multiplication quaternionique ($\mathrm{End}_{\overline{K}}(A)$ est un ordre dans une algèbre de quaternions sur $\mathbb{Q}$), et à multiplication complexe. Comme les variétés CM (sans restriction sur la dimension) feront l'objet du successif chapitre 5, nous nous restreignons ici aux trois premiers cas, pour lesquels nous donnons une description des groupes G_{ℓ^∞} pour ℓ suffisamment grand explicite. Notamment, pour le cas générique nous prouvons :

Théorème 17. *Soit A/K une surface abélienne telle que $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. Soit ℓ un premier qui n'est pas divisible par aucune place de mauvaise réduction de A. Si ℓ est non ramifié dans K et est strictement plus grand que $b(2 \cdot 1920[K : \mathbb{Q}], 4, 2h(A))^{1/4}$, alors $G_{\ell^\infty} = \mathrm{GSp}_4(\mathbb{Z}_\ell)$.*

Pour le cas de la multiplication réelle nous nous plaçons dans le cadre plus général des variétés de type GL_2, considérées par exemple par Ribet dans sa thèse [109] : pour des telles variétés nous montrons

Théorème 18. *Soit A/K une variété abélienne de dimension g telle que $\mathrm{End}_{\overline{K}}(A)$ est un ordre dans un corps totalement réel E de degré g sur $\mathbb{Q}$. Supposons que tous les endomorphismes de A sont définis sur K. Soit ℓ un nombre premier non ramifié dans $K \cdot E$ et strictement plus grand que $b(2[K : \mathbb{Q}], 2\dim(A), 2h(A))^{1/2}$ et $b(A/K)^g$. On a*

$$G_{\ell^\infty} = \left\{ x \in \mathrm{GL}_2\left(\mathcal{O}_E \otimes \mathbb{Z}_\ell\right) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}.$$

Finalement, nous avons un résultat du même type aussi pour le cas de la multiplication quaternionique :

Théorème 19. *Soit A/K une surface abélienne telle que $R = \mathrm{End}_{\overline{K}}(A)$ est un ordre dans une algèbre de quaternions (indéfinie) et soit Δ le discriminant de R. Supposons que tous les endomorphismes de A sont définis sur K. Soit ℓ un nombre premier strictement plus grand que $b(2[K : \mathbb{Q}], 4, 2h(A))^{1/2}$, qui ne divise pas Δ, et qui est non ramifié dans K. On a $G_{\ell^\infty} = (R \otimes \mathbb{Z}_\ell)^\times$.*

Faisons quelques remarques sur les outils impliqués dans la preuve de ces trois théorèmes concernant les surfaces abéliennes. Un des ingrédients essentiels pour la preuve du théorème 17 est une analyse des sousgroupes propres maximaux de $\mathrm{GSp}_4(\mathbb{F}_\ell)$. On peut grossièrement classifier ces sousgroupes en trois catégories :

- sous-groupes géométriques, qui préservent certaines structures additionnelles sur le $\mathbb{F}_\ell$-espace vectoriel $A[\ell]$: tombent dans cette catégorie par exemple les sous-groupes qui stabilisent une droite ou bien un plan de $\mathbb{F}_\ell^4 \cong A[\ell]$, aussi bien que les sous-groupes dont l'action sur $A[\ell]$ est semi-linéaire par rapport à une certaine structure de $\mathbb{F}_{\ell^2}$-espace vectoriel de $A[\ell]$;

- sous-groupes finis (dont l'image projective est de cardinal borné indépendamment de ℓ) : il s'agit de "petits" groupes finis qui admettent un plongement dans $\mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$ pour des familles infinies de premiers ℓ. Par exemple, certains groupes finis admettant une représentation symplectique irréductible de dimension 4 (sur $\mathbb{C}$) sont de ce type, car une telle représentation peut être réalisée sur $\mathbb{F}_\ell$ pour tout premier ℓ qui satisfait à certaines conditions de congruence ;
- sous-groupes dont l'image projective est conjuguée à l'image de la représentation symplectique $\mathrm{PGL}_2(\mathbb{F}_\ell) \to \mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$.

Pour montrer le résultat voulu il suffit de prouver que, si ℓ est suffisamment grand, l'image de Galois dans $\mathrm{Aut}\, A[\ell]$ n'est pas contenue dans aucun sous-groupe propre maximal de $\mathrm{GSp}_4(\mathbb{F}_\ell)$. Il s'agit alors de montrer que chacun des trois cas précédents ne peut pas se produire quand ℓ est suffisamment grand. On traite les cas géométriques grâce au théorème d'isogénie, que l'on applique à plusieurs réprises à la variété A, aussi bien qu'à certaines variétés abéliennes auxiliares construites à partir de A.
On peut également traiter le cas des petits groupes finis à l'aide du théorème d'isogénie, et c'est l'approche qu'on suit dans ce chapitre. Remarquons toutefois que dans le chapitre 4 nous développons une méthode plus fine, qui montre que ce cas ne peut pas se produire lorsque ℓ est plus grand qu'une certaine borne qui est essentiellement uniforme en la variété A. Finalement, le cas le plus difficile est celui des sous-groupes associés à la représentation $\mathrm{PGL}_2(\mathbb{F}_\ell) \to \mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$. Pour montrer que de tels sousgroupes ne peuvent pas contenir l'image de Galois, on applique des résultats profonds, dûs à Raynaud, qui décrivent l'action sur $A[\ell]$ du groupe d'inertie associé à une place de K de caractéristique ℓ. Le résultat souhaité est alors obtenu en comparant la description de cette action avec les propriétés de la représentation $\mathrm{PGL}_2(\mathbb{F}_\ell) \to \mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$.

Le chapitre 4 est consacré à l'étude effective des représentations associées aux variétés de dimension impaire. La preuve donnée par Serre de son théorème 4 ne se prête pas directement à être rendue effective, et nous sommes obligés de prendre un chemin assez différent. Après avoir rappelé une classification grossière des sous-groupes maximaux de $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$, nous exploitons les propriétés des représentations des groupes finis de type Lie pour montrer que si l'image de Galois est petite par rapport à $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$, alors en fait elle est contenue dans l'image d'une certaine représentation qui est *algébrique* (et définie en caractéristique 0). Nous pouvons alors disposer de toutes les techniques classiques de théorie des représentations, et nous prouvons que cette situation ne peut se produire que dans de cas très particuliers.
Cela ne nous conduit pas à une preuve effective du théorème 4 en toute dimension, mais nous obtenons quand même un résultat effectif quitte à imposer des restrictions additionnelles sur la dimension g. Le cas le plus favorable est celui de la dimension 3, pour lequel nous avons le résultat suivant :

Théorème 20. *Soit A/K une variété abélienne de dimension 3 telle que $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. Notons $\mathcal{N}^0_{A/K}$ le conducteur naif de A/K, à savoir le produit des idéaux premiers de $\mathcal{O}_K$ auxquels A a mauvaise réduction, et supposons que les points de 7-torsion de A sont tous définis sur K.*

- *Si l'Hypothèse de Riemann Généralisée est vraie, on a l'égalité* $G_{\ell^\infty} = \mathrm{GSp}_6(\mathbb{Z}_\ell)$ *pour tout premier* ℓ *non ramifié dans* K *et strictement plus grand que* $(2q)^{48}$*, où*
$$q = b(A^2/K;3)^8 \left(\log|\Delta_{K/\mathbb{Q}}| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right)^2.$$
- *Inconditionnellement, la même conclusion est vraie avec*
$$q = \exp\left(cb(A^2/K;3)^8 \left(\log|\Delta_K| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right)^2\right),$$
où c *est une constante absolue et effective.*

Nous obtenons également des résultats pour une infinité d'autres dimensions, mais la borne corréspondante n'est pas complètement explicite en fonction de la variété A.
Nous considérons finalement un exemple concret (la Jacobienne J d'une courbe hyperelliptique de genre 3 sur $\mathbb{Q}$) qui ne peut pas être traité par les méthodes existantes, et pour lequel nous déterminons explicitement une borne $b_0(J/\mathbb{Q})$ telle que l'action de Galois est maximale pour tout $\ell > b_0(J/\mathbb{Q})$.

Dans le chapitre 5 nous nous focalisons sur les variétés géométriquement simples admettant multiplication complexe. Nous établissons une version effective du théorème 10 pour N une puissance de nombre premier, et nous montrons notamment que dans ce cas les constantes C_1, C_2 y apparaissant peuvent être choisies indépendamment de A. Une version simplifiée du résultat principal de ce chapitre s'énonce ainsi :

Théorème 21. *Soit* K *un corps de nombres et* A/K *une variété abélienne de dimension* g*, admettant multiplication complexe sur* K *par un ordre d'un corps CM qu'on note* E*. Soit* r *le rang du groupe de Mumford-Tate de* A *et* ℓ *un premier plus grand que* $\sqrt{2\cdot g!}$ *et non ramifié dans* $E\cdot K$*. Soit* μ *le nombre de racines de l'unité en* E *et* $h(K)$ *le nombre de classe de* K*. On a l'encadrement :*
$$\frac{1}{4\mu\sqrt{g!}}\cdot \ell^{nr} \leq [K(A[\ell^n]) : K] \leq \frac{5}{2}\mu\cdot h(K)\cdot \ell^{nr}.$$

En fait, les résultats du chapitre 5 s'appliquent à tout nombre premier ℓ, mais les formules qu'on obtient sont moins simples à écrire. La preuve du théorème 21 répose sur plusieurs outils : d'abord bien sûr la théorie de la multiplication complexe, développée par Shimura-Taniyama et étendue par Weil, Serre et Tate, puis des arguments de théorie du corps de classes, et finalement une étude des isogénies entre tores définis sur des corps locaux, faite à la fois par des méthodes de cohomologie galoisienne et de géométrie différentielle p-adique.
Nous traitons avec plus de détail le cas des courbes elliptiques CM, pour lesquelles nous montrons le résultat adélique suivant, qui fournit une description très précise de l'image G_∞ :

Théorème 22. *Soit* E/K *une courbe elliptique telle que* $\mathrm{End}_{\overline{K}}(E)$ *est un ordre dans le corps quadratique imaginaire* F*. Notons* $\rho_\infty : \mathrm{Gal}(\overline{K}/K) \to \prod_\ell \mathrm{Aut}\, T_\ell E$ *la représentation adélique associée à* E *et* G_∞ *son image. Pour tout premier* ℓ *notons ensuite* C_ℓ *le groupe* $(\mathcal{O}_F\otimes\mathbb{Z}_\ell)^\times$*, considéré comme sous-groupe de* $\mathrm{Aut}_{\mathbb{Z}_\ell}(\mathcal{O}_F\otimes\mathbb{Z}_\ell) \cong \mathrm{GL}_2(\mathbb{Z}_\ell) \cong \mathrm{Aut}\, T_\ell E$*, et* $N(C_\ell)$ *le normalisateur de* C_ℓ *dans* $\mathrm{GL}_2(\mathbb{Z}_\ell)$*.*

1. *Supposons que $F \subseteq K$: alors G_∞ est inclus dans $\prod_\ell C_\ell$, et l'indice $[\prod_\ell C_\ell : G_\infty]$ est borné par $3[K : \mathbb{Q}]$. De plus, l'égalité $G_\ell = C_\ell$ est vérifiée pour tout ℓ non ramifié dans K et de bonne réduction pour E.*

2. *Supposons que $F \not\subseteq K$: alors G_∞ est inclus dans $\prod_\ell N(C_\ell)$ mais pas dans $\prod_\ell C_\ell$, et l'indice $[\prod_\ell N(C_\ell) : G_\infty]$ n'est pas fini. L'intersection $H_\infty = G_\infty \cap \prod_\ell C_\ell$ a indice 2 dans G_∞, et l'indice $[\prod_\ell C_\ell : H_\infty]$ est borné par $6[K : \mathbb{Q}]$. Finalement on a $G_\ell = N(C_\ell)$ pour tout ℓ non ramifié dans $K \cdot F$ et de bonne réduction pour E.*

Si de plus on suppose que l'invariant j de E n'est pas égal à 0 ni à 1728, alors les constantes 3 and 6 des parties (1) et (2) peuvent être remplacées par 1 et 2 respectivement.

Toujours dans le chapitre 5, nous entamons aussi une étude assez approfondie du groupe de Mumford-Tate des variétés CM, qui – comme il est bien connu – est un tore dans ce cas. La difficulté principale est ici celle de comprendre le groupe des $\mathbb{Z}_\ell$-points d'un $\mathbb{Q}$-tore ayant mauvaise réduction en ℓ : sous l'hypothèse que $\mathrm{MT}(A)$ soit de dimension $\dim A + 1$, ce qui est bien le cas générique pour une variété CM, nous pouvons effectuer des calculs locaux qui conduisent à des bornes essentiellement optimales pour les degrés $[K(A[\ell^n]) : K]$:

Théorème 23. *Soit A/K une variété abélienne géometriquement simple de dimension g, admettant multiplication complexe (sur $\overline{K}$) par le corps CM E. Soit $\mathrm{MT}(A)$ le groupe de Mumford-Tate de A et r son rang.*

1. *Si ℓ est non ramifié en E on a l'encadrement suivant :*
$$(1 - 1/\ell)^r \ell^{nr} \leq |\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \leq (1 + 1/\ell)^r \ell^{nr}.$$

2. *Supposons $r = g + 1$. Alors pour tout premier $\ell \neq 2$ et tout $n \geq 1$ on a*
$$(1 - 1/\ell)^{g+1} \cdot \ell^{(g+1)n} \leq |\,\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \leq 2^g\,(1 + 1/\ell)^{g-1}\,\ell^{(g+1)n},$$
et pour $\ell = 2$ et $n \geq 1$ on a
$$\frac{1}{2^{2g+3}} \cdot 2^{(g+1)n} \leq |\,\mathrm{MT}(A)(\mathbb{Z}/2^n\mathbb{Z})| \leq \frac{1}{2} 4^g \cdot 2^{(g+1)n}.$$

Ce résultat est particulièrement intéressant pour les variétés CM de dimension au plus 3, car toute telle variété respecte automatiquement l'hypothèse $\mathrm{rg}\,\mathrm{MT}(A) = \dim(A) + 1$.

Dans la deuxième partie de la thèse nous quittons le domaine de l'effectivité pour nous tourner vers des questions de nature plus qualitative.

Dans le chapitre 6 nous développons des techniques qui permettent d'étudier les groupes algébriques $\mathcal{G}_\ell(A \times B)$ associés à un produit $A \times B$ de deux variétés abéliennes, et nous donnons une condition suffisante pour que les groupes $\mathcal{G}_\ell(A), \mathcal{G}_\ell(B)$ déterminent le groupe $\mathcal{G}_\ell(A \times B)$. Nos résultats s'expriment plus aisément à l'aide de l'objet suivant :

Définition 24. Soit A/K une variété abélienne, ℓ un nombre premier, et $\mathcal{G}_\ell(A)$ le groupe de monodromie algébrique associé à A/K. Nous posons $\mathcal{H}_\ell(A) := \left(\mathcal{G}_\ell(A) \cap \mathrm{SL}_{T_\ell(A)\otimes\mathbb{Q}_\ell}\right)^0$.

Notre condition suffisante s'énonce alors ainsi (pour le cas d'un corps de définition de caractéristique zéro : nous traitons aussi le cas d'un corps de définition quelconque) :

Théorème 25. *Soit K un corps finiment engendré de caractéristique zéro, soient A_1 et A_2 deux variétés abéliennes sur K, et soit ℓ un nombre premier. Pour $i = 1, 2$ soit $\mathfrak{h}_i$ l'algèbre de Lie de $\mathcal{H}_\ell(A_i)$. Supposons les conditions suivantes vérifiées :*

1. *pour $i = 1, 2$ l'algèbre $\mathfrak{h}_i$ est semisimple (on a donc $\mathfrak{h}_i \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{h}_{i,1} \oplus \cdots \oplus \mathfrak{h}_{i,n_i}$, où chaque $\mathfrak{h}_{i,j}$ est simple) ;*
2. *pour $i = 1, 2$, il existe une décomposition $V_\ell(A_i) \otimes \overline{\mathbb{Q}_\ell} \cong V_{i,1} \oplus \cdots \oplus V_{i,n_i}$ telle que l'action de $\mathfrak{h}_i \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{h}_{i,1} \oplus \cdots \oplus \mathfrak{h}_{i,n_i}$ sur $V_{i,1} \oplus \cdots \oplus V_{i,n_i}$ se fait composante par composante, et $\mathfrak{h}_{i,j}$ agit de façon fidèle sur $V_{i,j}$;*
3. *pour tout choix de (i, j) et (i', j') distincts tels qu'il existe un isomorphisme $\varphi : \mathfrak{h}_{i,j} \to \mathfrak{h}_{i',j'}$, il existe une représentation irréductible W de $\mathfrak{h}_{i,j}$ telle que tous les sous-modules simples de $V_{i,j}$ et de $\varphi^*\left(V_{i',j'}\right)$ (considérés comme représentations de $\mathfrak{h}_{i,j}$) sont isomorphes à W, et le plus haut poids qui définit W est stable sous l'action de tous les automorphismes de $\mathfrak{h}_{i,j}$.*

Alors on a soit $\mathrm{Hom}_{\overline{K}}(A_1, A_2) \neq 0$, *soit* $\mathcal{H}_\ell(A_1 \times A_2) \cong \mathcal{H}_\ell(A_1) \times \mathcal{H}_\ell(A_2)$.

Même si les hypothèses de ce résultat sont assez techniques, elles sont tout de même très souvent vérifiées, et nous déduisons de ce théorème plusieurs critères facilement applicables qui donnent des conditions suffisantes pour que l'égalité $\mathcal{H}_\ell(A_1 \times A_2) \cong \mathcal{H}_\ell(A_1) \times \mathcal{H}_\ell(A_2)$ soit vérifiée.
Par exemple, nous montrons comme un résultat d'Ichikawa [44], prouvé initialement pour les structures de Hodge, peut être transposé au cadre galoisien :

Théorème 26. *Soit K un corps finiment engendré de caractéristique zéro et soient A'_i, A''_j (pour $i = 1, \ldots, n$ et $j = 1, \ldots, m$) des K-variétés abéliennes absolument simples, de dimension relative impaire, et deux à deux non isogènes sur $\overline{K}$. Supposons que chaque A'_i est de type I, II ou III au sens de la classification d'Albert, alors que chaque A''_j est de type IV. Soit finalement A une K-variété abélienne qui est $\overline{K}$-isogène à $\prod_{i=1}^n A'_i \times \prod_{j=1}^m A''_j$. On a alors*

$$\mathcal{H}_\ell(A) \cong \prod_{i=1}^{n} \mathcal{H}_\ell\left(A'_i\right) \times \mathcal{H}_\ell\left(\prod_{j=1}^{m} A''_j\right).$$

Nous appliquons ensuite ce théorème, et plus généralement les méthodes de ce chapitre, pour prouver que la conjecture de Mumford-Tate est vraie pour toute variété abélienne de dimension au plus 5 dont tous les facteurs géométriquement simples vérifient eux-mêmes Mumford-Tate. Cela complète un résultat précédent de Moonen et Zarhin [81], qui déterminent les classes de Hodge sur de telles variétés abéliennes non simples.

On étudie enfin une question soulevée naturellement par les travaux de Hindry et Ratazzi [38] [39] [40]. Dans leur étude des extensions engendrées par les sous-groupes de torsion d'une variété abélienne, les deux auteurs introduisent deux variantes ("forte" et "faible") d'une propriété qu'ils appellent (μ), et ils montrent que les deux variantes sont satisfaites par certaines (amples) classes de variétés abéliennes. Rappelons ici leur définitions. Si A/K est une variété abélienne, le choix d'une polarisation induit, pour tout ℓ et tout n, un accouplement (de Weil)

$$e_{\ell^n} : A[\ell^n] \times A[\ell^n] \to \mu_{\ell^n}.$$

Fixons maintenant un sous-groupe fini H de $A[\ell^\infty]$. Comme H est fini, il est en particulier d'exposant fini, d'où un entier m tel que $H \subseteq A[\ell^m]$. Comme l'accouplement de Weil est Galois-équivariant, on voit immédiatement que le groupe de racines de l'unité $e_{\ell^m}(H \times H)$ est contenu dans $K(H)$. Grossièrement, on dit que A respecte la propriété (μ) forte si toute racine de l'unité ℓ^m-ième dans $K(H)$ est dans l'image de $e_{\ell^m}(H \times H)$:

Définition 27. Soit A/K une variété abélienne (avec une K-polarisation fixée). Posons

$$m_1(H) = \max\left\{ k \;\middle|\; \begin{array}{c} \exists n \geq 0,\ \exists P, Q \in H \text{ d'ordre } \ell^n, \text{ tels que} \\ e_{\ell^n}(P,Q) \text{ engendre } \mu_{\ell^k} \end{array} \right\}.$$

On dit que A/K respecte la propriété μ forte par rapport à la polarisation fixée s'il existe une constante C (dépendant de A et K, mais pas de ℓ) telle que pour tout premier ℓ, et pour tout sous-groupe fini H de $A[\ell^\infty]$, on a l'encadrement

$$\frac{1}{C}\left[K\left(\mu_{\ell^{m_1(H)}}\right) : K\right] \leq [K(H) \cap K(\mu_{\ell^\infty})] \leq C\left[K\left(\mu_{\ell^{m_1(H)}}\right) : K\right].$$

Par ailleurs, il est facile de voir que cette définition ne dépend pas de la polarisation choisie : une variété abélienne respecte la propriété (μ) forte pour une certaine polarisation si et seulement si elle la respecte pour toute polarisation. On peut également considérer une version un peu affaiblie de cette propriété, à savoir

Définition 28. Soit A/K une variété abélienne (avec une K-polarisation fixée). On dit que A/K respecte la propriété μ faible par rapport à la polarisation fixée s'il existe une constante C (dépendant de A et K, mais pas de ℓ) telle que pour tout premier ℓ, et pour tout sousgroupe fini H de $A[\ell^\infty]$, on a l'encadrement

$$\frac{1}{C}\left[K\left(\mu_{\ell^n}\right) : K\right] \leq [K(H) \cap K(\mu_{\ell^\infty})] \leq C\left[K\left(\mu_{\ell^n}\right) : K\right]$$

pour un certain n qui peut dépendre de ℓ et de H.

L'intérêt de ces propriétés est lié, comme le montrent Hindry et Ratazzi, à l'étude asymptotique du nombre des points de torsion de $A(F)$, où A/K est une variété abélienne fixée et F parcourt les extensions finies de K. Plus précisement, ces conditions (μ) – forte et faible – apparaissent de manière naturelle lorsque l'on cherche à calculer l'invariant $\gamma(A/K)$ suivant :

$$\gamma(A/K) = \inf\left\{ x > 0 \;\middle|\; \begin{array}{c} \exists C \in \mathbb{R} :\ \text{pour toute extension finie } F/K \\ \text{on a } |A(F)_{\text{tors}}| \leq C[F:K]^x \end{array} \right\}.$$

Dans le chapitre 7 nous prouvons que la version faible de la propriété (μ) est vraie pour toute variété abélienne qui vérifie la conjecture de Mumford-Tate, et on montre que un célèbre exemple dû à Mumford [85] fournit un contre-exemple à la propriété (μ) forte dans le cas général, même si on suppose que notre variété abélienne a un anneau d'endomorphismes trivial.

Introduction

In this thesis we consider various classes of problems concerning Galois representations attached to abelian varieties, with a particular emphasis on the question of obtaining effective results. To fix the notation, our objects of interest are abelian varieties A defined over number fields K, and the associated families of Galois representations

$$\rho_{\ell^\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\, T_\ell A,$$

where $T_\ell(A)$ denotes as usual the ℓ-adic Tate module of A. We denote by G_{ℓ^∞} the image of ρ_{ℓ^∞} and by $\mathcal{G}_\ell$ the $\mathbb{Q}_\ell$-algebraic group obtained by taking the Zariski closure of G_{ℓ^∞} in $\mathrm{GL}_{T_\ell(A)\otimes\mathbb{Q}_\ell}$. We consider three different aspects of the study of the ρ_{ℓ^∞}'s:

1. determination of $\mathcal{G}_\ell$ for non-simple A.

2. horizontal situation: certain combinations of conditions on $\mathrm{End}_K(A)$ and $\dim A$ imply that G_{ℓ^∞} is of a specific form when ℓ is large enough; for example, when $\dim A$ is either 2 or an odd number and $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, then it is known by work of Serre [116] [118] [121] that $G_{\ell^\infty} = \mathrm{GSp}_{2\dim A}(\mathbb{Z}_\ell)$ for all ℓ large enough. For some such cases we quantify what "large enough" means in terms of A and K.

3. vertical situation: for a given A/K and a given prime ℓ consider the $\mathbb{Z}_\ell$-closure $\mathcal{G}_{\mathbb{Z}_\ell}$ of G_{ℓ^∞} in $\mathrm{GL}_{T_\ell A}$. The group G_{ℓ^∞} is of finite index in $\mathcal{G}_{\mathbb{Z}_\ell}(\mathbb{Z}_\ell)$, and in some cases we can give an explicit upper bound on the index $[\mathcal{G}_{\mathbb{Z}_\ell}(\mathbb{Z}_\ell) : G_{\ell^\infty}]$.

When A is an elliptic curve (with or without CM) both parts (2) and (3) can be carried out successfully, and we actually obtain *adelic* results, which take into account all representations ρ_{ℓ^∞} at the same time. Since we are chiefly concerned with the problem of giving *effective* descriptions of the groups G_{ℓ^∞}, what we want to do is bound the complexity of the representations G_{ℓ^∞} in terms of the arithmetic complexity of A: thanks to the work of Faltings [26], we dispose of a canonical measure of complexity for A, its semistable Faltings height $h(A)$, and it is in terms of $h(A)$ that we express our results. Works of Raynaud [105], Masser and Wüstholz [72] [70], and Gaudron and Rémond [28] make it possible to effectively describe the isogeny class of A, and we extract from this information finer data about the groups G_{ℓ^∞}.

Concerning the **determination of** $\mathcal{G}_\ell$ we prove that some numerical conditions on the dimensions of the abelian varieties A, B allow us to compute $\mathcal{G}_\ell(A \times B)$ in terms of $\mathcal{G}_\ell(A)$ and $\mathcal{G}_\ell(B)$, recovering in this context analogous results that were already known for the Hodge groups of non-simple abelian varieties. As an application, we show that our criterion implies the truth of the Mumford-Tate

conjecture for non-simple abelian varieties of dimension at most 5, assuming that each geometrically simple factor satisfies the conjecture (recall that Mumford-Tate is known to hold for absolutely simple abelian varieties of dimension at most 3 and for most absolutely simple abelian fourfolds). To show such statements we mainly use techniques issued from representation theory, as pioneered by Serre, Ribet [111][109] and Zarhin [152], combined with results of Larsen and Pink [59] [58] on the structure of the Lie algebra of $\mathcal{G}_\ell$.

A prototypical example of result in the **horizontal setting** is the following description of the groups G_{ℓ^∞} attached to abelian surfaces:

Theorem 1. *There is an explicit polynomial function $f(d,h)$ with the following property. Let A/K be an abelian surface with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, and let ℓ be a rational prime that is not divisible by any place of bad reduction of A and does not ramify in K. If $\ell > f([K:\mathbb{Q}], h(A))$, then $G_{\ell^\infty} = \mathrm{GSp}_4(\mathbb{Z}_\ell)$.*

We also prove similar results for (geometrically simple) surfaces with arbitrary endomorphism rings, for abelian varieties of GL_2-type, and for geometrically simple threefolds A with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$; we also establish weaker results in the same vein for some abelian varieties of higher (odd) dimension satisfying $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. The techniques involved are in a sense more sophisticated than the ones used to treat the algebraic groups $\mathcal{G}_\ell$: the main issue is the need to work with representations in positive characteristic, and a combination of classical representation theory in characteristic zero with the methods of finite group theory (especially Aschbacher's classification of maximal subgroups of classical finite groups) is necessary to treat such problems. Group-theoretical methods alone, however, do not suffice: on the more arithmetical side we also need to rely on Chebotarev's density theorem and on results of Raynaud [104] that describe the action of inertia on the Galois modules $A[\ell]$.

The **vertical situation** is the most delicate one. Zywina [156] was the first to obtain adelic results for non-CM elliptic curves, but his approach was limited to the field $\mathbb{Q}$, and his bound not completely effective; by contrast, in this thesis we obtain fully explicit results for (products of) non-CM elliptic curves over any number field. Specifically, for a single elliptic curve we prove the following version of Serre's celebrated open image theorem:

Theorem 2. *Let E/K be an elliptic curve that does not admit complex multiplication. Denote by*

$$\rho_\infty : \mathrm{Gal}(\overline{K}/K) \to \prod_\ell \mathrm{Aut}\, T_\ell(A) \cong \mathrm{GL}_2(\hat{\mathbb{Z}})$$

the adelic representation attached to E. The inequality

$$\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right] < C_1 \cdot [K:\mathbb{Q}]^{C_2} \cdot \max\{1, h(E), \log[K:\mathbb{Q}]\}^{2C_2}$$

holds, where $C_1 = \exp(6 \cdot 10^{29527})$ and $C_2 = 4.9 \cdot 10^{10}$.

By refining the same idea through a more in-depth study of the combinatorics of subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$ we are then able to extend this result to cover the case of arbitrary products of elliptic curves without complex multiplication. The main technical tool underlying these results is a construction, due to Pink [97], that allows for a classification of pro-ℓ subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ in terms of linear data, encoded in the form of certain Lie algebras with coefficients in $\mathbb{Z}_\ell$. However, Pink's

construction is not flexible enough for our purposes, and we need to extend his results in various directions, circumventing for example the difficulties arising from the fact that his approach cannot be made to work for the prime $\ell = 2$. Once our extended construction is in place for $\mathrm{GL}_2(\mathbb{Z}_\ell)$ and $\mathrm{GL}_2(\mathbb{Z}_\ell)^2$, the general case of an arbitrary product of copies of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ follows by applying an integral version of the Goursat-Ribet lemma.

The one case in which the vertical situation is relatively well-understood is that of abelian varieties admitting complex multiplication. Mainly thanks to work of Ribet [110], relying on previous fundamental contributions of Shimura-Taniyama and Serre-Tate, in the CM case estimates are available for the degrees over K of the division fields $K(A[\ell^n])$, but such estimates are not completely effective, and in fact it is not entirely clear what their dependence on A should be. Relying by and large on the approach of Ribet, but combining it with techniques issued from Galois cohomology (and even with some p-adic differential geometry), we can make such estimates effective and show in particular that they can be made essentially uniform in the variety, the only relevant information being the rank of the Mumford-Tate group of A. A rough but easily-stated version of this result reads as follows:

Theorem 3. *Let K be a number field and A/K be an abelian variety of dimension g admitting complex multiplication over K by an order in the CM field E. Denote by μ be the number of roots of unity contained in E and by $h(K)$ the class number of K. Let r be the rank of the Mumford-Tate group of A and $\ell > \sqrt{2 \cdot g!}$ be a prime unramified in $E \cdot K$. The following inequality holds:*

$$\frac{1}{4\mu\sqrt{g!}} \cdot \ell^{nr} \leq [K(A[\ell^n]) : K] \leq \frac{5}{2}\mu \cdot h(K) \cdot \ell^{nr}.$$

Finally, in a slightly different direction, exploiting again methods coming from the representation theory of algebraic groups (in positive characteristic), we explore a rather peculiar phenomenon which does not seem to have been noticed before. Specifically, if H is a finite subgroup of $A[\ell^\infty]$, we show that while the degree of $K(H) \cap K(\mu_{\ell^\infty})$ over K is essentially a power of ℓ (up to bounded factors), it is *not* true that this power of ℓ is determined by the knowledge of the image of the Weil pairing $H \times H \to \mu_{\ell^\infty}$. As shown by Hindry and Ratazzi [38] [39], this fact has a bearing on the study of the extensions of K generated by torsion points of A.

Chapter 1

Adelic bounds for representations arising from elliptic curves

1.1 Introduction

We are interested in studying Galois representations attached (via ℓ-adic Tate modules) to elliptic curves E defined over an arbitrary number field K and without complex multiplication, i.e. such that $\mathrm{End}_{\overline{K}}(E) = \mathbb{Z}$. Let us recall briefly the setting and fix some notation: the action of $\mathrm{Gal}(\overline{K}/K)$ on the torsion points of $E_{\overline{K}}$ gives rise to a family of representations (indexed by the rational primes ℓ)

$$\rho_\ell : \mathrm{Gal}\left(\overline{K}/K\right) \to \mathrm{GL}(T_\ell(E)),$$

where $T_\ell(E)$ denotes the ℓ-adic Tate module of E. As $T_\ell(E)$ is a free module of rank 2 over $\mathbb{Z}_\ell$ it is convenient to fix bases and regard these representations as morphisms

$$\rho_\ell : \mathrm{Gal}\left(\overline{K}/K\right) \to \mathrm{GL}_2(\mathbb{Z}_\ell),$$

and it is the image G_ℓ of these maps that we aim to study. It is also natural to encode all these representations in a single 'adelic' map

$$\rho_\infty : \mathrm{Gal}\left(\overline{K}/K\right) \to \mathrm{GL}_2(\widehat{\mathbb{Z}}),$$

whose components are the ρ_ℓ and whose image we denote G_∞. By a theorem of Serre ([116, §4, Théorème 3]) G_∞ is open in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$, and the purpose of the present study is to show that the adelic index $[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : G_\infty]$ is in fact bounded by an explicit function depending only on the stable Faltings height $h(E)$ of E and on the degree of K over $\mathbb{Q}$, generalizing and making completely explicit a result proved by Zywina [156] in the special case $K = \mathbb{Q}$. More precisely we show:

Theorem 1.1.1. *Let E/K be an elliptic curve that does not admit complex multiplication. The inequality*

$$\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right] < \gamma_1 \cdot [K:\mathbb{Q}]^{\gamma_2} \cdot \max\left\{1, h(E), \log[K:\mathbb{Q}]\right\}^{2\gamma_2}$$

holds, where $\gamma_1 = \exp(10^{21483})$ and $\gamma_2 = 2.4 \cdot 10^{10}$.

Remark 1.1.2. We actually prove a more precise result (theorem 1.9.1), from which the present bound follows through elementary estimates. The large constants appearing in this theorem have a

very strong dependence on those of theorem 1.2.1; unpublished results that Eric Gaudron and Gaël Rémond have been kind enough to share with the author show that the statement can be improved to

$$\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right] < \gamma_3 \cdot \left([K:\mathbb{Q}] \cdot \max\left\{1, h(E), \log[K:\mathbb{Q}]\right\}\right)^{\gamma_4}$$

with the much better constants $\gamma_3 = \exp\left(1.9 \cdot 10^{10}\right)$ and $\gamma_4 = 12395$, cf. remark 1.9.4.

As an easy corollary we also get:

Corollary 1.1.3. *Let E/K be an elliptic curve that does not admit complex multiplication. There exists a constant $\gamma(E/K)$ with the following property: for every $x \in E_{\mathrm{tors}}(\overline{K})$ (of order denoted $N(x)$) the inequality*

$$[K(x):K] \geq \gamma(E/K) N(x)^2$$

holds. We can take $\gamma(E/K) = \left(\zeta(2) \cdot \left[\,\mathrm{GL}_2\left(\widehat{\mathbb{Z}}\right) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\,\right]\right)^{-1}$, which can be explicitly bounded thanks to the main theorem.

Remark 1.1.4. This corollary (with the same proof, but with a non-effective $\gamma(E/K)$) follows directly from the aforementioned theorem of Serre ([116, §4, Théorème 3]). The exponent 2 for $N(x)$ is best possible, as is easily seen from the proof by taking $N = \ell$, a prime large enough that $G_\ell = \mathrm{GL}_2(\mathbb{Z}_\ell)$. It should also be pointed out that for a general (possibly CM) elliptic curve Masser ([67, p. 262]) proves an inequality of the form

$$[K(x):K] \geq \gamma'(K) h(E)^{-3/2} \frac{N(x)}{\log N(x)},$$

where $\gamma'(K)$ is an effectively computable (but non-explicit) constant that only depends on $[K:\mathbb{Q}]$.

We briefly sketch the proof strategy, highlighting differences and similarities between our approach and that of [156]. By a technique due to Masser and Wüstholz (cf. [71], [72] and [68]), and which is by now standard, it is possible to give a bound on the largest prime ℓ for which the representation modulo ℓ is not surjective; an argument of Serre then shows that (for $\ell \geq 5$) this implies full ℓ-adic surjectivity. This eliminates all the primes larger than a computable bound (actually, of all those that do not divide a quantity that can be bounded explicitly in terms of E). We then have to deal with the case of non-surjective reduction, that is, with a finite number of 'small' primes.
In [156] these small primes are treated using two different techniques. All but a finite number of them are dealt with by studying a family of Lie algebras attached to G_ℓ; this analysis is greatly simplified by the fact that the reduction modulo ℓ of G_ℓ is not contained in a Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$, a result depending on the hard theorem of Mazur on cyclic ℓ-isogenies. The remaining primes belong to an explicit list (again given by Mazur's results), and are treated by an application of Faltings' theorem to certain modular curves. This approach, however, has two important drawbacks. On the one hand, effective results on cyclic isogenies do not seem – at present – to be available for arbitrary number fields, so the use of Mazur's theorem is a severe obstacle in generalizing this technique to number fields larger than $\mathbb{Q}$. On the other hand, and perhaps more importantly, the use of Faltings' theorem is a major hindrance to effectivity, since making the result explicit for a given number field K would require understanding the K-points of a very large number of modular curves, a task that currently seems to be far beyond our reach.

While we do not introduce any new ideas in the treatment of the large primes, relying by and large on the methods of Masser-Wüstholz, we do put forward a different approach for the small primes that allows us to bypass both the difficulties mentioned above. With respect to [156], the price to pay to avoid the use of Mazur's theorem is a more involved analysis of the Lie algebras associated with subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$, which is done here without using a congruence filtration, but dealing instead with all the orders at the same time; this approach seems to be more natural, and proves more suitable for generalization to arbitrary number fields. We also avoid the use of Faltings' theorem entirely. This too comes at a cost, namely replacing uniform bounds with functions of the Faltings height of the elliptic curve, but it has the advantage of giving a completely explicit result, which does not depend on the (potentially very complicated) arithmetic of the K-rational points on the modular curves.

The organization of the present chapter reflects the steps alluded to above: in section 1.2 we recall an explicit form of the isogeny theorem (as proved by Gaudron and Rémond in [28] building on the work of Masser and Wüstholz) and an idea of Masser that will help improve many of the subsequent estimates by replacing an inequality with a divisibility condition. In sections 3 through 6 we prove the necessary results on the relation between Lie algebras and closed subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$; the main technical tool we use to show that the Galois image is large is the following theorem, which is proved in sections 1.4 (for odd ℓ) and 1.5 (for $\ell = 2$):

Theorem 1.1.5. *Let ℓ be an odd prime (resp. $\ell = 2$). For every closed subgroup G of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (resp. every closed subgroup whose reduction modulo 2 is trivial if $\ell = 2$) define $L(G)$ to be the $\mathbb{Z}_\ell$-span of $\left\{g - \frac{\mathrm{tr}(g)}{2} \cdot \mathrm{Id} \;\middle|\; g \in G\right\}$.*
Let H be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$. There is a closed subgroup H_1 of H, of index at most 24 (resp. with trivial reduction modulo 2 and of index at most 192 for $\ell = 2$), such that the following implication holds for all positive integers s: if $L(H_1)$ contains $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$, then H_1 itself contains

$$\mathcal{B}_\ell(4s) = \left\{g \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; g \equiv \mathrm{Id} \pmod{\ell^{4s}}\right\} \quad \textit{(resp. } \mathcal{B}_2(6s) \textit{ for } \ell = 2\textit{)}.$$

The methods of these sections are then applied in section 1.7 to get bounds valid for *every* prime ℓ (cf. theorem 1.7.5, which might have some independent interest), while section 1.8 deals with the large primes through the aforementioned ideas of Masser and Wüstholz. Finally, in section 1.9 we put it all together to get the adelic estimate.

1.2 Preliminaries on isogeny bounds

The main tool that makes all the effective estimates possible is a very explicit isogeny-type theorem taken from [28], which builds on the seminal work of Masser and Wüstholz (cf. [70] and [72]). To state it we will need some notation: we let $\alpha(g) = 2^{10}g^3$ and define, for any abelian variety A/K of dimension g,

$$b([K:\mathbb{Q}], g, h(A)) = \left((14g)^{64g^2}[K:\mathbb{Q}]\max\left(h(A), \log[K:\mathbb{Q}], 1\right)^2\right)^{\alpha(g)}.$$

Theorem 1.2.1. *([28] Théorème 1.4; cf. also the section 'Cas elliptique') Let K be a number field and A, A^* be two abelian K-varieties of dimension g. If A, A^* are isogenous over K, then there*

exists a K-isogeny $A^* \to A$ whose degree is bounded by $b([K:\mathbb{Q}], \dim(A), h(A))$. If E is an elliptic curve without complex multiplication over $\overline{K}$, then the same holds with $b([K:\mathbb{Q}], \dim(A), h(A))$ replaced by

$$10^{13}[K:\mathbb{Q}]^2 \max\left(h(E), \log[K:\mathbb{Q}], 1\right)^2.$$

Remark 1.2.2. As the notation suggests, the three arguments of b will always be the degree of a number field K, the dimension g of an abelian variety A/K and its stable Faltings height $h(A)$.

Remark 1.2.3. Unpublished results of Gaudron and Rémond show that if A is the N-th power of an elliptic curve E/K and A^* is K-isogenous to A, then a K-isogeny $A^* \to A$ exists whose degree does not exceed $10^{13N}[K:\mathbb{Q}]^{2N} \max\left(h(E), \log[K:\mathbb{Q}], 1\right)^{2N}$.

The following theorem follows easily from the arguments in Masser's paper [68]; however, since it is never stated explicitly in the form we need, in the interest of completeness we include a short proof.

Theorem 1.2.4. *(Masser) Suppose that A/K is an abelian variety that is isomorphic over K to a product $A_1^{e_1} \times \ldots \times A_n^{e_n}$, where $A_1, \ldots, A_n$ are simple over K, mutually non-isogenous over K, and have trivial endomorphism ring over K. Let $b \in \mathbb{R}$ be a constant with the following property: for every K-abelian variety A^* isogenous to A over K there exists an isogeny $\psi : A^* \to A$ with $\deg \psi \leq b$. Then there exists an integer $b_0 \leq b$ with the following property: for every K-abelian variety A^* isogenous to A over K there exists an isogeny $\psi_0 : A^* \to A$ with $\deg \psi_0 \mid b_0$.*

Proof. We take the notation of [68], which we briefly recall. Let m be a positive integer and G be a $\mathrm{Gal}(\overline{K}/K)$-submodule of $A[m]$. For every K-endomorphism τ of A we denote by $\ker_m \tau$ the intersection $\ker \tau \cap A[m]$; we also define

$$f_m(G) := \min_{\tau} \left[\ker_m \tau : G\right],$$

where the minimum is taken over all τ in $\mathrm{End}_K(A)$ with $G \subseteq \ker_m \tau$. By [68, Lemma 3.3] we have $f_m(G) \leq b$ for every positive integer m and every Galois submodule G of $A[m]$. We set $b_0 := \max_{m,G} f_m(G)$, where the maximum is taken over all positive integers m and all Galois submodules G of $A[m]$: clearly we have $b_0 \leq b$. Now if A^* is a K-abelian variety that is K-isogenous to A over K, then by [68, Lemma 4.1] there exists a K-isogeny $\psi : A^* \to A$ such that $\deg \psi \mid b_0$, and this establishes the theorem. Notice that in order to apply [68, Lemma 4.1] we need $i(\mathrm{End}_K(A)) = 1$ (in the notation of [68]), which can be deduced as in [68, p. 185, proof of Theorem 2]. □

We will denote by $b_0(K, A)$ the minimal b_0 with the property of the above theorem; in particular $b_0(K, A) \leq b([K:\mathbb{Q}], h(A), \dim(A))$. Consider now $b_0(K', A)$ as K' ranges over the finite extensions of K of degree bounded by d. On one hand, $b_0(K, A)$ divides $b_0(K', A)$; on the other hand $b_0(K', A) \leq b(d[K:\mathbb{Q}], h(A), \dim(A))$ stays bounded, and therefore the number

$$b_0(K, A; d) = \operatorname*{lcm}_{[K':K] \leq d} b_0(K', A)$$

is finite. The function $b_0(K, A; d)$ is studied in [68], Theorem D, mostly through the following elementary lemma:

Lemma 1.2.5. *([68, Lemma 7.1]) Let $X, Y \geq 1$ be real numbers and $\mathcal{B}$ be a family of natural numbers. Suppose that for every positive integer t and every subset A of $\mathcal{B}$ with $|A| = t$ we have $\mathrm{lcm}(A) \leq XY^t$. The least common multiple of the elements of $\mathcal{B}$ is then finite, and does not exceed $4^{eY}X^{1+\log(C)}$, where $e = \exp(1)$.*

By adapting Masser's argument to the function $b(d[K:\mathbb{Q}], h(A), \dim(A))$ at our disposal it is immediate to prove:

Proposition 1.2.6. *If A is of dimension $g \geq 1$ and satisfies the hypotheses of the previous theorem, then*

$$b_0(K, A; d) \leq 4^{\exp(1)\cdot\left(d(1+\log d)^2\right)^{\alpha(g)}} b([K:\mathbb{Q}], \dim(A), h(A))^{1+\alpha(g)(\log(d)+2\log(1+\log d))}.$$

If E is an elliptic curve without complex multiplication over $\overline{K}$, then the number $b_0(K, E; d)$ is bounded by

$$4^{\exp(1)\cdot d^2(1+\log d)^2}\left(10^{13}[K:\mathbb{Q}]^2 \max\left(h(E), \log[K:\mathbb{Q}], 1\right)^2\right)^{1+2\log d+2\log(1+\log d)}.$$

Proof. We can clearly assume $d \geq 2$. We apply the lemma to $\mathcal{B} = \{b_0(K', A)\}_{[K':K]\leq d}$. Choose t elements of $\mathcal{B}$, corresponding to extensions $K_1, \ldots, K_t$ of K, and set $L = K_1 \cdots K_t$. We claim that

$$\max\left\{\log(d^t[K:\mathbb{Q}]), 1\right\} \leq (1+\log(d))^t \max\left\{1, \log[K:\mathbb{Q}]\right\}.$$

Indeed the right hand side is clearly at least 1, so it suffices to show the inequality

$$t\log(d) + \log[K:\mathbb{Q}] \leq (1+\log(d))^t \max\left\{1, \log[K:\mathbb{Q}]\right\};$$

as $\log(d) > 0$, we have $(1+\log(d))^t \geq 1 + t\log(d)$ by Bernoulli's inequality, and the claim follows. We thus see that $\mathrm{lcm}(b_0(K_1, A), \ldots, b_0(K_t, A))$ divides

$$\begin{aligned} b_0(L, A) &\leq b([L:\mathbb{Q}], \dim(A), h(A)) \\ &\leq b(d^t[K:\mathbb{Q}], \dim(A), h(A)) \\ &\leq \left(\left(d(1+\log d)^2\right)^{\alpha(g)}\right)^t b([K:\mathbb{Q}], \dim(A), h(A)), \end{aligned}$$

so we can apply the above lemma with

$$X = b([K:\mathbb{Q}], \dim(A), h(A)),\ Y = \left(d(1+\log d)^2\right)^{\alpha(g)}$$

to get the desired conclusion. The second statement is proved in the same way using the corresponding improved bound for elliptic curves. □

Remark 1.2.7. We are only going to use the function $b_0(K, A; d)$ for bounded values of d (in fact, $d \leq 24$), so the essential feature of the previous proposition is to show that, under this constraint, $b_0(K, A; d)$ is bounded by a polynomial in $b([K:\mathbb{Q}], \dim(A), h(A))$.
Also notice that, if $A = E^2$ is the square of an elliptic curve E/K, then using the improved version of theorem 1.2.1 mentioned in remark 1.2.3 we get

$$b_0\left(K, E^2; d\right) \leq 4^{\exp(1)\cdot d^4(1+\log d)^4}\left(10^{26}[K:\mathbb{Q}]^4 \max\left(h(E), \log[K:\mathbb{Q}], 1\right)^4\right)^{1+4\log d+4\log(1+\log d)}.$$

We record all these facts together as a theorem for later use:

Theorem 1.2.8. *Suppose A/K is an abelian variety, isomorphic over K to a product of simple abelian varieties, each having trivial endomorphism ring over K. There exists a positive integer*

$b_0(K,A)$, not exceeding $b([K:\mathbb{Q}],\dim(A),h(A))$, with the following property: if A^* is isogenous to A over K, then there exists an isogeny $A^* \to A$, defined over K, whose degree divides $b_0(K,A)$. Furthermore, for every fixed d the function

$$b_0(K,A;d) = \mathrm{lcm}_{[K':K]\leq d}\, b_0(K',A)$$

exists and is bounded by a polynomial in $b([K:\mathbb{Q}],\dim(A),h(A))$.

1.3 Group theory for $\mathrm{GL}_2(\mathbb{Z}_\ell)$

Let ℓ be any rational prime. The subject of the following four sections is the study of certain Lie algebras associated with closed subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$; the construction we present is inspired by Pink's paper [97], but we will have to extend his results in various directions: in particular, our statements apply to $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (and not just to $\mathrm{SL}_2(\mathbb{Z}_\ell)$), to *any* ℓ, including 2, and to arbitrary (not necessarily pro-ℓ) subgroups. The present section contains a few necessary, although elementary, preliminaries on congruence subgroups, and introduces the relevant objects and notations.

1.3.1 Congruence subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)$

We aim to study the structure of the congruence subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)$, which we denote

$$\mathcal{B}_\ell(n) = \left\{ x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; x \equiv \mathrm{Id} \pmod{\ell^n} \right\}.$$

Notation. We let v_ℓ be the standard discrete valuation of $\mathbb{Z}_\ell$ and set $v = v_\ell(2)$ (namely $v = 0$ if $\ell \neq 2$ and $v = 1$ otherwise). We also let $\binom{\frac{1}{2}}{k}$ denote the generalized binomial coefficient $\binom{\frac{1}{2}}{k} = \frac{1}{k!}\prod_{i=0}^{k-1}\left(\frac{1}{2}-i\right)$ and define $\sqrt{1+t}$ to be the formal power series $\sum_{k\geq 0}\binom{\frac{1}{2}}{k}t^k$.

The first piece of information we need is the following description of a generating set for $\mathcal{B}_\ell(n)$:

Lemma 1.3.1. *For $n \geq 1$ the group $\mathcal{B}_\ell(n)$ is generated by the elements*

$$L_a = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix},\ R_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \text{ and } D_c = \begin{pmatrix} 1+c & 0 \\ 0 & \frac{1}{1+c} \end{pmatrix}$$

for a, b, c ranging over $\ell^n\mathbb{Z}_\ell$.

Proof. Let $x = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ be an element of $\mathcal{B}_\ell(n)$. Since $x_{11} \equiv 1 \pmod{\ell}$, it is in particular a unit, so $a = -\dfrac{x_{21}}{x_{11}}$ has valuation $v_\ell(a) = v_\ell(x_{21}) \geq n$, i.e. $a \in \ell^n\mathbb{Z}_\ell$. Next we compute

$$L_a x = \begin{pmatrix} x_{11} & x_{12} \\ 0 & ax_{12}+x_{22} \end{pmatrix};$$

we are thus reduced to the case $x_{21} = 0$. Under this hypothesis, and choosing $b = -\dfrac{x_{12}}{x_{11}}$, it is easily seen that $xR_b \in \mathcal{B}_\ell(n)$ is diagonal, and since every diagonal matrix in $\mathcal{B}_\ell(n)$ is by definition of the form D_c for some $c \in \ell^n\mathbb{Z}_\ell$ we are done. □

We will also need a description of the derived subgroup of $\mathcal{B}_\ell(n)$; in order to prove the relevant result, we first need a simple-minded lemma on valuations that will actually come in handy in many instances:

Lemma 1.3.2. *Let $x \in \mathbb{Z}_\ell$. We have:*

1. *For $\ell = 2$ and $v_2(x) \geq 3$ the series $\sqrt{1+x} = \sum_{k\geq 0} \binom{\frac{1}{2}}{k} x^k$ converges to the only solution λ of the equation $\lambda^2 = 1+x$ that satisfies $\lambda \equiv 1 \pmod 4$. The inequality $v_2(\sqrt{1+x}-1) \geq v_2(x)-1$ holds.*
2. *For $\ell \neq 2$ and $v_\ell(x) > 0$ the series $\sqrt{1+x} = \sum_{k\geq 0} \binom{\frac{1}{2}}{k} x^k$ converges to the only solution λ of the equation $\lambda^2 = 1+x$ that satisfies $\lambda \equiv 1 \pmod \ell$. The equality $v_\ell(\sqrt{1+x}-1) = v_\ell(x)$ holds.*

Proof. For $\ell = 2$ we have

$$v_2\left(\binom{\frac{1}{2}}{k}\right) = v_2\left(\frac{(1/2)(-1/2)...(-(2k-3)/2)}{k!}\right) = -k - v_2(k!) \geq -2k,$$

while for any other prime

$$v_\ell\left(\binom{\frac{1}{2}}{k}\right) = v_\ell\left(\prod_{i=1}^{k-1}(2i-1)\right) - v_\ell(k!) \geq -v_\ell(k!) \geq -\frac{1}{\ell-1}k.$$

Convergence of the series is then immediate in both cases, and the identity of power series

$$\left(\sum_{k\geq 0}\binom{\frac{1}{2}}{k}t^k\right)^2 = 1+t$$

implies that, for every x such that the series converges, $\sum_{k\geq 0}\binom{\frac{1}{2}}{k}x^k$ is indeed a solution to the equation $\lambda^2 = 1+x$.

Let now $\ell = 2$. Note that in the series expansion $\sqrt{1+x}-1 = \sum_{k\geq 1}\binom{\frac{1}{2}}{k}x^k$ all the terms, except perhaps the first one, have valuation at least

$$(v_2(x)-2)\cdot 2 \geq v_2(x) - 1;$$

as for the first term, it is simply $\frac{x}{2}$, so it has exact valuation $v_2(x)-1$ and we are done; a similar argument works for $\ell \neq 2$, except now $v_\ell\left(\frac{x}{2}\right) = v_\ell(x)$. The congruence $\sqrt{1+x} \equiv 1 \pmod 4$ (resp. modulo ℓ) now follows. □

Lemma 1.3.3. *For $n \geq 1$ the derived subgroup of $\mathcal{B}_\ell(n)$ contains $\mathcal{B}_\ell(2n+2v)$.*

Proof. Take $R_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ with $b \equiv 0 \pmod{\ell^{2n+2v}}$ and set $\beta = \ell^n$. By the above lemma $1+\frac{b}{\beta}$ has a square root y congruent to 1 modulo ℓ that automatically satisfies $y \equiv 1 \pmod{\ell^n}$, so

$$M = \begin{pmatrix} y & 0 \\ 0 & \frac{1}{y} \end{pmatrix} \text{ and } N = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}$$

both belong to $\mathcal{B}_\ell(n)$. It is immediate to compute

$$MNM^{-1}N^{-1} = \begin{pmatrix} 1 & \beta(y^2-1) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix},$$

so R_b is an element of $\mathcal{B}_\ell(n)'$. Similar identities also show that, for every $a \equiv 0 \pmod{2^{2n+2v}}$, the derived subgroup $\mathcal{B}_\ell(n)'$ contains $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} = L_a$. To finish the proof (using lemma 1.3.1) we now just need to show that $\mathcal{B}_\ell(n)'$ contains D_c for every $c \equiv 0 \pmod{\ell^{2n+2v}}$. This is done through an identity similar to the above, namely we set

$$M = \begin{pmatrix} \sqrt{1+c} & 0 \\ \frac{-c}{\beta\sqrt{1+c}} & \frac{1}{\sqrt{1+c}} \end{pmatrix} \text{ and } N = \begin{pmatrix} 1 & \beta \\ \frac{c}{\beta} & c+1 \end{pmatrix}$$

and compute that $MNM^{-1}N^{-1} = \begin{pmatrix} 1+c & 0 \\ 0 & \frac{1}{1+c} \end{pmatrix} = D_c$. The only thing left to check is that M and N actually belong to $\mathcal{B}_\ell(n)$, which is easily done by observing that $\sqrt{1+c} \equiv 1 \pmod{\ell^n}$ by the series expansion and that $v_\ell\left(\frac{-c}{\beta\sqrt{1+c}}\right) \geq 2n+2v-n \geq n$. □

To conclude this paragraph we describe a finite set of generators for the congruence subgroups of $\mathrm{SL}_2(\mathbb{Z}_2)$:

Lemma 1.3.4. *Let $a, u \in \mathbb{Z}_2$ and $L_a = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$. Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$. If $L_a \in G$, then G also contains $L_{au} = \begin{pmatrix} 1 & 0 \\ au & 1 \end{pmatrix}$. Similarly, if G contains $R_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$, then it also contains R_{bu} for every $u \in \mathbb{Z}_2$. Finally, if $c \equiv 0 \pmod 4$ and G contains $D_c = \begin{pmatrix} 1+c & 0 \\ 0 & \frac{1}{1+c} \end{pmatrix}$, then G contains D_{cu} for every $u \in \mathbb{Z}_2$.*
Let s be an integer no less than 2. If $a, b, c \in 4\mathbb{Z}_2$ are such that $\max\{v_2(a), v_2(b), v_2(c)\} \leq s$, and if G contains L_a, R_b and D_c, then G contains $\mathcal{B}_2(s)$.

Proof. We show that the set W consisting of the w in $\mathbb{Z}_2$ such that L_{aw} belongs to G is a closed subgroup of $\mathbb{Z}_2$ containing 1. Indeed, $L_{aw_1}L_{aw_2} = L_{a(w_1+w_2)}$ by an immediate direct calculation, so in particular $L_{aw}^{-1} = L_{-aw}$; furthermore $1 \in W$ by hypothesis, and if w_n is a sequence of elements of W converging to w, then $\{L_{aw_n}\} \subseteq G$ converges to L_{aw}, and since G is closed L_{aw} itself belongs to G, so $w \in W$. It follows that W is closed and contains the integers, and since $\mathbb{Z}$ is dense in $\mathbb{Z}_2$ we get $W = \mathbb{Z}_2$ as claimed. Given that $u \mapsto R_{bu}$ is a group morphism the same proof also works for the family R_{bu}. The situation with the family D_{cu} is slightly different, in that $u \mapsto D_{cu}$ is not a group morphism; however, if $w \in \mathbb{Z}_2$, then we see that

$$(D_c)^w = \begin{pmatrix} (1+c)^w & 0 \\ 0 & \frac{1}{(1+c)^w} \end{pmatrix}$$

is well-defined and belongs to G (indeed this is trivially true for $w \in \mathbb{Z}$, and then we just need argue by continuity). As $c \equiv 0 \pmod 4$ we also have the identity $(1+c)^w = \exp(w\log(1+c))$, since all the involved power series converge: more precisely, for any γ in $4\mathbb{Z}_2$ the series $\sum_{j=1}^{\infty}(-1)^{j+1}\frac{\gamma^j}{j}$

converges and defines $\log(1+\gamma)$, and since the inequality $v_2(\gamma^j) - v_2(j) > v_2(\gamma)$ holds for every $j \geq 2$ we have $v_2(\log(1+\gamma)) = v_2(\gamma) \geq 2$. Suppose now that $v_2(\gamma) \geq v_2(c)$: then $w = \frac{\log(1+\gamma)}{\log(1+c)}$ exists in $\mathbb{Z}_2$, so we can consider $(1+c)^w = \exp(w\log(1+c)) = \exp(\log(1+\gamma)) = 1+\gamma$ and therefore for any such γ the matrix D_γ belongs to G. The last statement is now an immediate consequence of lemma 1.3.1. □

1.3.2 Lie algebras attached to subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$

Our study of the groups G_ℓ will go through suitable integral Lie algebras, for which we introduce the following definition:

Definition 1.3.5. Let A be a commutative ring. A **Lie algebra over** A is a finitely presented A-module M together with a bracket $[\cdot,\cdot] : M \times M \to M$ that is A-bilinear, antisymmetric and satisfies the Jacobi identity. For any A, the module $\mathfrak{sl}_2(A) = \left\{M \in M_2(A) \;\middle|\; \mathrm{tr}(M) = 0\right\}$ endowed with the usual commutator is a Lie algebra over A. The same is true for $\mathfrak{gl}_2(A)$, the set of all 2×2 matrices with coefficients in A.

We restrict our attention to the case $A = \mathbb{Z}_\ell$, and try to understand closed subgroups G of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ by means of a surrogate of the usual Lie algebra construction. In order to do so, we introduce the following definitions, inspired by those of [97]:

Definition 1.3.6. Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$; if $\ell = 2$, suppose that the image of G in $\mathrm{GL}_2(\mathbb{F}_2)$ is trivial. We set

$$\begin{array}{rccc} \Theta: & G & \to & \mathfrak{sl}_2(\mathbb{Z}_\ell) \\ & g & \mapsto & g - \frac{1}{2}\mathrm{tr}(g)\cdot \mathrm{Id}\,. \end{array}$$

Note that this definition makes sense even for $\ell = 2$, since by hypothesis the 2-adic valuation of the trace of g is at least 1.

Definition 1.3.7. The **special Lie algebra** of G, denoted $L(G)$ (or simply L if no confusion can arise), is the closed subgroup of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ topologically generated by $\Theta(G)$. We further define $C(G)$, or simply C, as the closed subgroup of $\mathbb{Z}_\ell$ topologically generated by all the traces $\mathrm{tr}(xy)$ for x, y in $L(G)$.

Remark 1.3.8.

1. $L(G)$ is indeed a Lie algebra because of the identity
$$[\Theta(x), \Theta(y)] = \Theta(xy) - \Theta(yx).$$
2. If G is a subgroup of H then $L(G)$ is contained in $L(H)$.
3. C is a $\mathbb{Z}_\ell$-module: indeed it is a $\mathbb{Z}$-module, and the action of $\mathbb{Z}$ is continuous for the ℓ-adic topology, so it extends to an action of $\mathbb{Z}_\ell$ since C is closed. Therefore C is an ideal of $\mathbb{Z}_\ell$.

The key importance of $L(G)$, at least for odd ℓ, lies in the following result:

Theorem 1.3.9. *([97, Theorem 3.3]) Let ℓ be an odd prime and G be a pro-ℓ subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$. Set $L_2 = [L(G), L(G)]$ and*

$$H_2 = \left\{x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; \Theta(x) \in L_2, \mathrm{tr}(x) - 2 \in C(G)\right\}.$$

Then H_2 is the derived subgroup of G.

On the other hand, for $\ell = 2$ the property of Θ that will be crucial for our study of L is the following approximate addition formula:

Lemma 1.3.10. *([97, Formula 1.3]) For every $g_1, g_2 \in \mathrm{GL}_2(\mathbb{Z}_\ell)$, if $\ell \neq 2$ (respectively for every $g_1, g_2 \in \left\{x \in \mathrm{GL}_2(\mathbb{Z}_2) \;\middle|\; \mathrm{tr}(x) \equiv 0 \pmod 2\right\}$, for $\ell = 2$), the following identity holds:*

$$2\left(\Theta(g_1 g_2) - \Theta(g_1) - \Theta(g_2)\right) = [\Theta(g_1), \Theta(g_2)] + (\mathrm{tr}(g_1) - 2)\,\Theta(g_2) + (\mathrm{tr}(g_2) - 2)\,\Theta(g_1).$$

In what follows we will often want to recover partial information on G from information about the reduction of G modulo various powers of ℓ. It is thus convenient to use the following notation:

Notation. We denote by $G(\ell^n)$ the image of the reduction map $G \to \mathrm{GL}_2(\mathbb{Z}/\ell^n\mathbb{Z})$. We also let π be the projection map $G \to G(\ell)$.

We now record a simple fact about modules over DVRs we will need later:

Lemma 1.3.11. *Let A be a DVR, n a positive integer, M a subset of A^n and $N = \langle M \rangle$ the submodule of A^n generated by M. Denote by π_k the projection $A^n \to A$ on the k-th component. There exist a basis $x_1, \dots, x_m$ of N consisting of elements of M and scalars $(\sigma_{ij})_{1 \leq j < i \leq m} \subseteq A$ with the following property: if we define inductively $t_1 = x_1$ and $t_i = x_i - \Sigma_{j<i}\sigma_{ij}t_j$ for $i \geq 2$, then $\pi_k\left(x_i - \Sigma_{j<l}\sigma_{ij}t_j\right) = 0$ for every $1 \leq k < l \leq i \leq m$. The t_j are again a basis of N.*

Proof. We proceed by induction on n. The case $n = 1$ is easy: M is just a subset of A, and the claim is that the ideal generated by M can also be generated by a single element of M, which is clear. Consider now a subset M of A^{n+1}. Let ν be the discrete valuation of A; the set $\left\{\nu(\pi_1(x)) \;\middle|\; x \in M\right\}$ consists of non-negative integers, therefore it admits a minimum k_1. Take x_1 to be any element of M such that $\nu(\pi_1(x_1)) = k_1$. For every element $m \in M$ we can form $f(m) = m - \dfrac{\pi_1(m)}{\pi_1(x_1)}x_1$, which is again an element of A^{n+1} since by definition of x_1 we have $\pi_1(x_1) \mid \pi_1(m)$. It is clear enough that $\pi_1(f(m)) = 0$ for all $m \in M$. Therefore $f(M)$ is a subset of $\{0\} \oplus A^n$, and it is also apparent that the module generated by x_1 and $f(M)$ is again N. Apply the induction hypothesis to $f(M)$ (thought of as a subset of A^n). It yields a basis $f(x_2), \dots, f(x_m)$ of $f(M)$, scalars $(\tau_{ij})_{2 \leq j < i \leq m}$, and a sequence $u_2 = f(x_2), u_i = f(x_i) - \sum_{2 \leq j < i} \tau_{ij}u_j$, such that $\pi_k(f(x_i) - \Sigma_{2 \leq j < l}\tau_{ij}u_j) = 0$ for $2 \leq k < l \leq i \leq m$. We also have $\pi_1(f(x_i) - \Sigma_{2 \leq j < l}\tau_{ij}u_j) = 0$ if we view the u_i as elements of A^{n+1}. It is now enough to show that, with this choice of the x_i, it is possible to find scalars $\sigma_{ij}, 1 \leq j < i \leq m$, in such a way that $t_i = u_i$ for $i \geq 2$, and this we prove again by induction. By definition $u_2 = f(x_2) = x_2 - \dfrac{\pi_1(x_2)}{\pi_1(x_1)}x_1$, so we can take $\sigma_{21} = \dfrac{\pi_1(x_2)}{\pi_1(x_1)}$. Assuming we have proved the result up to level i, then, we have

$$u_{i+1} = f(x_{i+1}) - \sum_{2 \leq j < i+1} \tau_{ij}u_j = x_{i+1} - \frac{\pi_1(x_{i+1})}{\pi_1(x_1)}x_1 - \sum_{2 \leq j < i+1} \tau_{ij}t_j,$$

and we simply need to take $\sigma_{i+1,1} = \dfrac{\pi_1(x_{i+1})}{\pi_1(x_1)}$ and $\sigma_{ij} = \tau_{ij}$.
As for the last statement, observe that the matrix giving the transformation from the x_i to the t_j is unitriangular, hence invertible. □

1.3.3 Subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell), \mathrm{SL}_2(\mathbb{Z}_\ell)$, and their reduction modulo ℓ

In view of the next sections it is convenient to recall some well-known facts about the subgroups of $\mathrm{GL}_2(\mathbb{F}_\ell)$, starting with the following definition:

Definition 1.3.12. A subgroup J of $\mathrm{GL}_2(\mathbb{F}_\ell)$ is said to be:

- **split Cartan**, if J is conjugated to the subgroup of diagonal matrices. In this case the order of J is prime to ℓ.
- **nonsplit Cartan**, if there exists a subalgebra A of $\mathrm{M}_2(\mathbb{F}_\ell)$ that is a field and such that $J = A^\times$. The order of J is prime to ℓ, and J is conjugated to $\left\{ \begin{pmatrix} a & b\varepsilon \\ b & a \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_\ell) \right\}$, where ε is a fixed quadratic nonresidue.
- **the normalizer of a split (resp. nonsplit) Cartan**, if there exists a split (resp. nonsplit) Cartan subgroup $\mathcal{C}$ such that J is the normalizer of $\mathcal{C}$. The index $[J : \mathcal{C}]$ is 2, and ℓ does not divide the order of J (unless $\ell = 2$).
- **Borel**, if J is conjugated to the subgroup of upper-triangular matrices. In this case J has a unique ℓ-Sylow, consisting of the matrices of the form $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$.
- **exceptional**, if the projective image $\mathbb{P}J$ of J in $\mathrm{PGL}_2(\mathbb{F}_\ell)$ is isomorphic to either A_4, S_4 or A_5, in which case the order of $\mathbb{P}J$ is either 12, 24 or 60.

The above classes essentially exhaust all the subgroups of $\mathrm{GL}_2(\mathbb{F}_\ell)$. More precisely we have:

Theorem 1.3.13. *(Dickson's classification, cf. [116]) Let ℓ be a prime number and J be a subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$. Then we have:*

- *if ℓ divides the order of J, then either J contains $\mathrm{SL}_2(\mathbb{F}_\ell)$ or it is contained in a Borel subgroup;*
- *if ℓ does not divide the order of J, then J is contained in a (split or nonsplit) Cartan subgroup, in the normalizer of one, or in an exceptional group.*

As subgroups of $\mathrm{SL}_2(\mathbb{F}_\ell)$ are in particular subgroups of $\mathrm{GL}_2(\mathbb{F}_\ell)$, the above classification also covers all subgroups of $\mathrm{SL}_2(\mathbb{F}_\ell)$. Cartan subgroups of $\mathrm{SL}_2(\mathbb{F}_\ell)$ are cyclic (both in the split and nonsplit case).

The next lemma can be proved by direct inspection of the group structure of A_4, S_4 and A_5, and will help us quantify how far exceptional subgroups are from being abelian:

Lemma 1.3.14. *The groups A_4 and S_4 have abelian subgroups of order N if and only if $1 \leq N \leq 4$. The group A_5 has abelian subgroups of order N if and only if $1 \leq N \leq 5$.*

The following lemma, due to Serre, will prove extremely useful in showing that $G_\ell = \mathrm{GL}_2(\mathbb{Z}_\ell)$ using only information about the reduction of G_ℓ modulo ℓ:

Lemma 1.3.15. *Let $\ell \geq 5$ be a prime and G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$. Suppose that the image of G in $\mathrm{SL}_2(\mathbb{F}_\ell)$ is equal to $\mathrm{SL}_2(\mathbb{F}_\ell)$: then $G = \mathrm{SL}_2(\mathbb{Z}_\ell)$. Similarly, if H is a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ whose image in $\mathrm{GL}_2(\mathbb{F}_\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$, then $H' = \mathrm{SL}_2(\mathbb{Z}_\ell)$.*

Proof. The first statement is [119, IV-23, Lemma 3]. For the second, consider the closed subgroup H' of $\mathrm{SL}_2(\mathbb{Z}_\ell)$. Since by assumption we have $\ell > 3$, the finite group $\mathrm{SL}_2(\mathbb{F}_\ell)$ is perfect, so the image of H' in $\mathrm{SL}_2(\mathbb{F}_\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)' = \mathrm{SL}_2(\mathbb{F}_\ell)$. It then follows from the first part of the lemma that $H' = \mathrm{SL}_2(\mathbb{Z}_\ell)$ as claimed. □

The following definition will prove useful to translate statements about subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ into analogous results for subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ and vice versa:

Definition 1.3.16. Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (resp. $\mathrm{GL}_2(\mathbb{F}_\ell)$). The **saturation** of G, denoted $\mathrm{Sat}(G)$, is the group generated in $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (resp. $\mathrm{GL}_2(\mathbb{F}_\ell)$) by G and $\mathbb{Z}_\ell^\times \cdot \mathrm{Id}$ (resp. $\mathbb{F}_\ell^\times \cdot \mathrm{Id}$). The group G is said to be **saturated** if $G = \mathrm{Sat}(G)$. We also denote by $G^{\det=1}$ the group $G \cap \mathrm{SL}_2(\mathbb{Z}_\ell)$ (resp. $G \cap \mathrm{SL}_2(\mathbb{F}_\ell)$).

Lemma 1.3.17. *The following hold:*

1. *For every closed subgroup G of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ the groups G and $\mathrm{Sat}(G)$ have the same derived subgroup and the same special Lie algebra.*

2. *The two associations $G \mapsto G^{\det=1}$ and $H \mapsto \mathrm{Sat}(H)$ are mutually inverse bijections between the sets*
$$\mathcal{G} = \left\{ G \text{ subgroup of } \mathrm{GL}_2(\mathbb{Z}_\ell) \;\middle|\; \begin{array}{c} G \text{ is saturated,} \\ \det(g) \text{ is a square for every } g \text{ in } G \end{array} \right\}$$
and
$$\mathcal{H} = \left\{ H \text{ subgroup of } \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; -\mathrm{Id} \in H \right\}.$$
For every G in $\mathcal{G}$ the groups G and $G^{\det=1}$ have the same derived subgroup and the same special Lie algebra.

3. *The map $G \mapsto \mathrm{Sat}(G)$ commutes with reducing modulo ℓ, i.e.*
$$(\mathrm{Sat}(G))(\ell) = \mathrm{Sat}(G(\ell)).$$
If ℓ is odd and G is saturated we also have $G(\ell)^{\det=1} = G^{\det=1}(\ell)$.

Proof. 1. The statement is obvious for the derived subgroup. As for the special Lie algebra, let λg be any element of $\mathrm{Sat}(G)$, where $\lambda \in \mathbb{Z}_\ell^\times$ and $g \in G$. As $L(G)$ is a $\mathbb{Z}_\ell$-module, $\Theta(\lambda g) = \lambda\Theta(g)$ belongs to $L(G)$, hence $L(\mathrm{Sat}(G)) \subseteq L(G)$. The other inclusion is trivial.

2. The first statement is immediate to check since the determinant of any homothety is a square; the other follows by writing $G = \mathrm{Sat}(H)$ and applying (1) to $(\mathrm{Sat}(H))^{\det=1} = H$ and $\mathrm{Sat}(H)$.

3. This is clear for the saturation. For $G \mapsto G^{\det=1}$ note that $G(\ell)^{\det=1}$ contains $G^{\det=1}(\ell)$, so we need to show the opposite inclusion. Take any matrix $[g]$ in $G(\ell)^{\det=1}$. By definition $[g]$ is the reduction of a certain $g \in G$ whose determinant is 1 modulo ℓ. As ℓ is odd and

$\det(g)$ is congruent to 1 modulo ℓ we can apply lemma 1.3.2 and write $\det(g) = \lambda^2$, where $\lambda = \sqrt{1 + (\det(g) - 1)}$ is congruent to 1 modulo ℓ. As G is saturated, it contains $\lambda^{-1}\,\mathrm{Id}$, hence also $\lambda^{-1}g$, whose determinant is 1 by construction. Furthermore, as $\lambda \equiv 1 \pmod{\ell}$, the two matrices $\lambda^{-1}g$ and g are congruent modulo ℓ. We have thus found an element of G of determinant 1 that maps to $[g]$, so $G^{\det=1} \to G(\ell)^{\det=1}$ is surjective.

□

Finally, since we will be mainly concerned with the pro-ℓ part of our groups, we will find it useful to give this object a name:

Notation. If G is a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ we write $N(G)$ for its maximal normal subgroup that is a pro-ℓ group.

The following lemma shows that $N(G)$ is well-defined and gives a description of it:

Lemma 1.3.18. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ and $\pi : G \to G(\ell)$ the projection modulo ℓ: then G admits a unique maximal normal pro-ℓ subgroup $N(G)$, which can be described as follows.*

1. *If $G(\ell)$ is of order prime to ℓ, then $N(G) = \ker \pi$ and $G(\ell) \cong \dfrac{G}{N(G)}$.*
2. *If the order of $G(\ell)$ is divisible by ℓ, and furthermore $G(\ell)$ is contained in a Borel subgroup, then $N(G)$ is the inverse image in G of the unique ℓ-Sylow S of $G(\ell)$.*
3. *If $G(\ell)$ is all of $\mathrm{SL}_2(\mathbb{F}_\ell)$, then $N(G) = \ker \pi$ and $G(\ell) \cong \dfrac{G}{N(G)}$.*

Proof. Let N be a pro-ℓ normal subgroup of G. The image $\pi(N)$ is a normal pro-ℓ subgroup of $G(\ell)$, hence it is trivial in cases (1) and (3) and it is either trivial or the unique ℓ-Sylow of $G(\ell)$ in case (2). In cases (1) and (3) it follows that $N \subseteq \ker \pi$, and since $\ker \pi$ is pro-ℓ we see that $\ker \pi$ is the unique maximal normal pro-ℓ subgroup of G. In case (2), let S be the unique ℓ-Sylow of $G(\ell)$. It is clear that N is contained in $\pi^{-1}(S)$, which on the other hand is pro-ℓ and normal in G. Indeed, by choosing an appropriate (triangular) basis for $G(\ell)$ we can define

$$\begin{array}{ccccc} G & \to & G(\ell) & \to & \mathbb{F}_\ell^\times \\ g & \mapsto & \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} & \mapsto & a, \end{array}$$

whose kernel is exactly $\pi^{-1}(S)$. □

1.4 Recovering G from $L(G)$, when ℓ is odd

Our purpose in this section (for $\ell \neq 2$) and the next (for $\ell = 2$) is to prove results that yield information on G from analogous information on $L(G)$. The statements we are aiming for are the following:

Theorem 1.4.1. *Let ℓ be an odd prime and G a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$.*

1. *Suppose that $G(\ell)$ is contained in a Cartan or Borel subgroup, and that $|G/N(G)| \neq 4$. Then the following implication holds for all positive integers s:*

$(\star)$ *if* $L(G)$ *contains* $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$*, then* $L(N(G))$ *contains* $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$*.*

2. *Without any assumption on* G*, there is a closed subgroup* H *of* G *that satisfies* $[G : H] \leq 12$ *and the conditions in (1) (so* H *has property* $(\star)$*).*

Theorem 1.4.2. *Let* ℓ *be an odd prime, and* G *a closed subgroup of* $\mathrm{GL}_2(\mathbb{Z}_\ell)$*.*

1. *Suppose that* G *satisfies the two conditions:*

   (a) $\det(g)$ *is a square in* $\mathbb{Z}_\ell^\times$ *for every* $g \in G$*;*

   (b) $\mathrm{Sat}(G)^{\det=1}$ *satisfies the hypotheses of theorem 1.4.1 (1).*

   Then the following implication holds for all positive integers s*:*

   $(\star\star)$ *if* $L(G)$ *contains* $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$*, then* G' *contains* $\mathcal{B}_\ell(4s)$*.*

2. *Without any assumption on* G*, either* $G' = \mathrm{SL}_2(\mathbb{Z}_\ell)$ *or there is a closed subgroup* H *of* G *that satisfies both* $[G : H] \leq 24$ *and the conditions in (1) (so* H *has property* $(\star\star)$*).*

Remark 1.4.3. Condition (b) can be made more explicit using the description of the maximal normal pro-ℓ subgroup given in lemma 1.3.18. The conditions on G can translated into conditions on $(\mathrm{Sat}(G))^{\det=1}(\ell)$: this group should be cyclic or have order divisible by ℓ and be contained in a Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$. In the first case we require $|(\mathrm{Sat}(G))^{\det=1}(\ell)| \neq 4$; in the second case we need $\left|\mathrm{Sat}(G)^{\det=1}(\ell)/S\right| \neq 4$, where S is the unique ℓ-Sylow of $\mathrm{Sat}(G)^{\det=1}(\ell)$. With this description, it is clear that condition (b) is true if $\mathrm{Sat}(G)^{\det=1}(\ell)$ contained in a Borel or Cartan subgroup and has order not divisible by 4.

Let us remark that the statements numbered (2) in the above theorems require a case by case analysis, which will be carried out in section 1.4.6 for theorem 1.4.2 (the proof of theorem 1.4.1 (2) is perfectly analogous). In the same section we will also show that part (1) of theorem 1.4.2 can be reduced to the corresponding statement in theorem 1.4.1, so the core of the problem lies in proving the result for $\mathrm{SL}_2(\mathbb{Z}_\ell)$. Before delving into the details of the proof (that involves a certain amount of calculations) we describe the general idea, which is on the contrary quite simple. The following paragraph should only be considered as outlining the main ideas, without any pretense of formality. If G is as in theorem 1.4.1 (1), then $G/N(G)$ is cyclic, and we can fix a generator $[g] \in G/N(G)$ that lifts to a certain $g \in G$. Denote by φ the operator $x \mapsto g^{-1}xg$: then φ acts on G and, since it fixes Id, also on $L(G)$. Furthermore it preserves $L(N(G)) \subseteq L(G)$ by normality of $N(G)$ in G, and obviously it fixes $\Theta(g)$. If we were working over $\mathbb{Q}_\ell$ instead of $\mathbb{Z}_\ell$ we would have a decomposition $L(G) \cong \langle\Theta(g)\rangle \oplus M$, where M is a φ-stable subspace of dimension 2, and the projection operator $p : L(G) \to M$ could be expressed as a polynomial in φ. We would also expect M to consist of elements coming from $N(G)$, because $\langle\Theta(g)\rangle$ is simply the special Lie algebra of $\langle g\rangle$; this would provide us with many nontrivial elements in $L(N(G))$. We would finally deduce the equality $L(N(G)) = \mathfrak{sl}_2(\mathbb{Q}_\ell)$ by exploiting the fact that $L(N(G))$ is a Lie algebra of dimension at least 2 that is also stable under φ. This point of view also suggests that we cannot expect the

theorem to hold when $G(\ell)$ is exceptional: if $G/N(G)$ is a simple group, then we expect the special Lie algebra of G not to be solvable, and since the only non-solvable subalgebra of $\mathfrak{sl}_2(\mathbb{Q}_\ell)$ is $\mathfrak{sl}_2(\mathbb{Q}_\ell)$ itself, $L(G)$ should be very large even if $N(G)$ is very small.
In what follows we prove (1) of theorem 1.4.1 first when $|G/N(G)| = 2$ and then in case $G(\ell)$ is respectively contained in a split Cartan, Borel, or nonsplit Cartan subgroup; we then discuss the optimality of the statement, showing through examples that it cannot be extended to the exceptional case and that ℓ^{2s} cannot be replaced by anything smaller. Finally, in section 1.4.6 we finish the proof of theorem 1.4.2.

Notation. For $x \in L(G)$ we set $\pi_{ij}(x) = x_{ij}$, the coefficient in the i-th row and j-th column of the matrix representation of x in $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. The maps π_{ij} are obviously linear and continuous.

1.4.1 The case $|G/N(G)| = 2$

Suppose first that $G(\ell)$ is contained in a Cartan subgroup, so that $G/N(G) \cong G(\ell)$. The only nontrivial element x in $G(\ell)$ satisfies the relations $x^2 = \mathrm{Id}$ and $\det(x) = 1$, so it must be $-\mathrm{Id}$. It follows that G contains an element g of the form $-\mathrm{Id} + \ell A$ for a certain $A \in \mathrm{M}_2(\mathbb{Z}_\ell)$. Considering the sequence

$$g^{\ell^n} = (-\mathrm{Id} + \ell A)^{\ell^n} = -\mathrm{Id} + O(\ell^{n+1})$$

and given that G is closed we see that $-\mathrm{Id}$ is in G. Next observe that for every $h \in G$ either h or $-h$ belongs to $N(G)$. If g_1, g_2, g_3 are elements of G such that $\Theta(g_1), \Theta(g_2), \Theta(g_3)$ is a basis for $L(G)$, then on the one hand for each i either g_i or $-g_i$ belongs to $N(G)$, and on the other hand $\Theta(-g_i) = -\Theta(g_i)$, so $L(G) = L(N(G))$ and the claim follows.
Next suppose $G(\ell)$ is contained in a Borel subgroup. We can assume that the order of $G(\ell)$ is divisible by ℓ, for otherwise $G(\ell)$ is cyclic and we are back in the previous case. The canonical projection $G \to G/N(G)$ factors as

$$\begin{array}{ccccc} G & \to & G(\ell) & \to & \mathbb{F}_\ell^\times \\ g & \mapsto & \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} & \mapsto & a, \end{array}$$

so if $G/N(G)$ has order 2 we can find in $G(\ell)$ an element of the form $\begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix}$. Taking the ℓ-th power of this element shows that $G(\ell)$ contains $-\mathrm{Id}$ and we conclude as above.

1.4.2 The split Cartan case

Suppose that $G(\ell)$ is contained in a split Cartan, so that, by choosing a suitable basis, we can assume that $G(\ell)$ is contained in the subgroup of diagonal matrices of $\mathrm{SL}_2(\mathbb{F}_\ell)$. Fix an element $g \in G$ such that $[g] \in G(\ell)$ is a generator. By assumption the order of $[g]$ is not 4, and by the previous paragraph we can assume it is not 2; furthermore it is not divisible by ℓ. The minimal polynomial of $[g]$ is then separable, and $[g]$ has two distinct eigenvalues in $\mathbb{F}_\ell^\times$. It follows that g can be diagonalized over $\mathbb{Z}_\ell$ (its characteristic polynomial splits by Hensel's lemma), and we can choose a basis in which $g = \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix}$, where a is an ℓ-adic unit. Note that our assumption that $|G(\ell)|$

does not divide 4 implies in particular that $a^4 \not\equiv 1 \pmod{\ell}$. A fortiori ℓ does not divide a^2-1, so the diagonal coefficients of $\Theta(g) = \begin{pmatrix} \frac{a^2-1}{2a} & 0 \\ 0 & -\frac{a^2-1}{2a} \end{pmatrix}$ are ℓ-adic units. The following lemma allows us to choose a basis of $L(G)$ containing $\Theta(g)$:

Lemma 1.4.4. *Suppose $g \in G$ is such that $\Theta(g)$ is not zero modulo ℓ. The algebra $L(G)$ admits a basis of the form $\Theta(g), \Theta(g_2), \Theta(g_3)$, where g_2, g_3 are in G.*

Proof. Recall that $L(G)$ is of rank 3 since it contains $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Start by choosing $g_1, g_2, g_3 \in G$ such that $\Theta(g_1), \Theta(g_2), \Theta(g_3)$ is a basis for $L(G)$. As $\Theta(g)$ is not zero modulo ℓ, from an equality of the form

$$\Theta(g) = \sum_{i=1}^{3} \lambda_i \Theta(g_i)$$

we deduce that at least one of the λ_i is an ℓ-adic unit, and we can assume without loss of generality that it is λ_1. But then

$$\Theta(g_1) = \lambda_1^{-1} \left(\Theta(g) - \lambda_2\Theta(g_2) - \lambda_3\Theta(g_3)\right),$$

and we can replace g_1 with g. □

Recall that we denote by φ the endomorphism of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ given by $x \mapsto g^{-1}xg$. We now prove that $L(N(G))$ is φ-stable and, more generally, describe the φ-stable subalgebras of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

Lemma 1.4.5. *Let ℓ be an odd prime, G a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$, N a normal closed subgroup of G and g an element of G. The special Lie algebra $L(N)$ is stable under φ.*

Proof. As $\Theta(N)$ generates $L(N)$ it is enough to prove that φ stabilizes $\Theta(N)$. Let $x = \Theta(n)$ for a certain $n \in N$: then

$$g^{-1}xg = g^{-1}\left(n - \frac{\mathrm{tr}(n)}{2}\,\mathrm{Id}\right)g = g^{-1}ng - \frac{\mathrm{tr}(g^{-1}ng)}{2}\,\mathrm{Id} = \Theta(g^{-1}ng),$$

and this last element is in $\Theta(N)$ since N is normal in G. □

Lemma 1.4.6. *Let s be a non-negative integer. Let L be a φ-stable Lie subalgebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ and x_{11}, x_{12}, x_{21}, y_{11}, y_{12}, y_{21} be elements of $\mathbb{Z}_\ell$ with $v_\ell(x_{21}) \leq s$ and $v_\ell(y_{12}) \leq s$. If L contains both $l_1 = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{pmatrix}$ and $l_2 = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & -y_{11} \end{pmatrix}$, then it contains all of $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.*

Proof. Consider first the case $x_{12} = y_{21} = 0$. We compute

$$\varphi(l_1) = \begin{pmatrix} x_{11} & 0 \\ a^2x_{21} & -x_{11} \end{pmatrix},$$

so L contains $\begin{pmatrix} x_{11} & 0 \\ a^2x_{21} & -x_{11} \end{pmatrix} - l_1 = \begin{pmatrix} 0 & 0 \\ (a^2-1)x_{21} & 0 \end{pmatrix}$, where by our hypothesis on a the valuation of the bottom-left coefficient is at most s. Analogously, L contains $\begin{pmatrix} 0 & (a^2-1)y_{12} \\ 0 & 0 \end{pmatrix}$, and since it is a Lie algebra it also contains the commutator

$$\left[\begin{pmatrix} 0 & (a^2-1)y_{12} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ (a^2-1)x_{21} & 0 \end{pmatrix}\right] = \begin{pmatrix} (a^2-1)^2x_{21}y_{12} & 0 \\ 0 & -(a^2-1)^2x_{21}y_{12} \end{pmatrix},$$

whose diagonal coefficients have valuation at most $2s$. This establishes the lemma in case x_{12} and y_{21} are both zero, since the three elements we have found generate $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$. The general case is then reduced to the previous one by replacing l_1, l_2 with

$$a^2\varphi(l_1) - l_1 = \begin{pmatrix} (a^2-1)x_{11} & 0 \\ (a^4-1)x_{21} & -(a^2-1)x_{11} \end{pmatrix}$$

and $a^{-2}\varphi(l_2) - l_2$, and noticing that since by assumption $\ell \nmid a^4 - 1$ we have $v_\ell((a^4-1)x_{21}) = v_\ell(x_{21})$ and $v_\ell((a^{-4}-1)y_{12}) = v_\ell(y_{12})$. □

We know from lemma 1.4.5 that $L(N(G))$ is φ-stable, so in order to apply lemma 1.4.6 to $L(N(G))$ we just need to find two elements l_1, l_2 in $L(N(G))$ with the property that $v_\ell \circ \pi_{21}(l_1) \leq s$ and $v_\ell \circ \pi_{12}(l_2) \leq s$. Since the values of the diagonal coefficients do not matter for the application of this lemma we will simply write $*$ for any diagonal coefficient appearing from now on. In particular we write $g_2, g_3, \Theta(g_2), \Theta(g_3)$ in coordinates as follows:

$$g_i = \begin{pmatrix} * & g_{12}^{(i)} \\ g_{21}^{(i)} & * \end{pmatrix}, \Theta(g_i) = \begin{pmatrix} * & g_{12}^{(i)} \\ g_{21}^{(i)} & * \end{pmatrix}.$$

As $[g]$ generates $G(\ell)$, for $i = 2, 3$ there exist $k_i \in \mathbb{N}$ such that $[g_i] = [g]^{k_i}$, or equivalently such that $g^{-k_i} g_i \in N(G)$. Since $\Theta(g), \Theta(g_2), \Theta(g_3)$ generate $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$, but the off-diagonal coefficients of $\Theta(g)$ vanish, we can choose two indices $i_1, i_2 \in \{2, 3\}$ such that $v_\ell \circ \pi_{21}(\Theta(g_{i_1})) \leq s$ and $v_\ell \circ \pi_{12}(\Theta(g_{i_2})) \leq s$. On the other hand, $L(N(G))$ contains

$$\Theta(g^{-k_i} g_i) = \Theta\left(\begin{pmatrix} a^{-k_i} & 0 \\ 0 & a^{k_i} \end{pmatrix}\begin{pmatrix} * & g_{12}^{(i)} \\ g_{21}^{(i)} & * \end{pmatrix}\right) = \begin{pmatrix} * & a^{-k_i} g_{12}^{(i)} \\ a^{k_i} g_{21}^{(i)} & * \end{pmatrix},$$

where $a^{\pm k_i}$ is an ℓ-adic unit. The ℓ-adic valuation of the off-diagonal coefficients of $\Theta(g^{-k_i} g_i)$ is then the same as that of the corresponding coefficients of $\Theta(g_i)$, and we find two elements $l_1 = \Theta(g^{-k_{i_1}} g_{i_1})$ and $l_2 = \Theta(g^{-k_{i_2}} g_{i_2})$ that satisfy $v_\ell \circ \pi_{21}(l_1) \leq s$ and $v_\ell \circ \pi_{12}(l_2) \leq s$ as required. We can now apply lemma 1.4.6 with $(L, g, l_1, l_2) = (L(N(G)), g, \Theta(g_{i_1}), \Theta(g_{i_2}))$ and deduce that $L(N(G))$ contains $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, as claimed.

1.4.3 The Borel case

Suppose $G(\ell)$ is included in a Borel subgroup. If the order of $G(\ell)$ is prime to ℓ, then $G(\ell)$ is in fact contained in a split Cartan subgroup, and we are reduced to the previous case. We can therefore assume without loss of generality that the order of $G(\ell)$ is divisible by ℓ. In this case we know that $N(G)$ is the inverse image in G of the unique ℓ-Sylow of $G(\ell)$, and that the canonical projection $G \to G/N(G)$ factors as

$$\begin{array}{ccccc} G & \to & G(\ell) & \to & \mathbb{F}_\ell^\times \\ g & \mapsto & \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} & \mapsto & a. \end{array}$$

Let H be the image of this map. The group H is cyclic and we can assume that its order does not divide 4: it is not 4 by hypothesis and if it is 1 or 2 we are done. Let g be any inverse image in G of a generator of H. The matrix representing g can be diagonalized over $\mathbb{Z}_\ell$ since the characteristic polynomial of $[g] \in G(\ell)$ is separable, and the same exact argument as in the previous paragraph

shows that we can choose a basis of $L(G)$ of the form $\Theta(g), \Theta(g_2), \Theta(g_3)$. By definition of H we see that for $i = 2, 3$ there is an integer k_i such that $[g_i] = [g]^{k_i}$ in $G/N(G)$, and the rest of the proof is identical to that of the previous paragraph.

1.4.4 The nonsplit Cartan case

Suppose now that $G(\ell)$ is contained in a nonsplit Cartan subgroup. Fix a $g \in G$ such that $[g]$ generates $G(\ell)$. We know that $[g]$ is of the form $\begin{pmatrix} [a] & [b\varepsilon] \\ [b] & [a] \end{pmatrix}$, where $[\varepsilon]$ is a fixed quadratic nonresidue modulo ℓ. In order to put g into a standard form we need the following elementary lemma, which is an ℓ-adic analogue of the Jordan canonical form over the reals.

Lemma 1.4.7. *Up to a choice of basis of $\mathbb{Z}_\ell^2$, the matrix representing g can be chosen to be of the form $\begin{pmatrix} a & b\varepsilon \\ b & a \end{pmatrix}$ for certain a, b, ε lifting $[a], [b], [\varepsilon]$, and where moreover a, b are ℓ-adic units.*

Proof. The characteristic polynomial of $[g]$ splits over $\mathbb{F}_\ell\left[\sqrt{[\varepsilon]}\right]$, so by Hensel's lemma the characteristic polynomial of g splits over $\mathbb{Z}_\ell[\sqrt{\varepsilon}]$. The two eigenvalues of g in $\mathbb{Z}_\ell[\sqrt{\varepsilon}]$ are of the form $a \pm b\sqrt{\varepsilon}$ for certain $a, b \in \mathbb{Z}_\ell$ (the notation is coherent: since the eigenvalues of $[g]$ are simply the projections of the eigenvalues of g, the elements a, b map to $[a], [b]$ modulo ℓ, respectively).
By definition of eigenvalue we can find a vector $\mathbf{v}_+ \in \mathbb{Z}_\ell[\sqrt{\varepsilon}]^2$ such that $g\mathbf{v}_+ = (a + b\sqrt{\varepsilon})\mathbf{v}_+$. Normalize $\mathbf{v}_+$ in such a way that at least one of its coordinates is an ℓ-adic unit, write $\mathbf{v}_+ = \mathbf{w} + \mathbf{z}\sqrt{\varepsilon}$ for certain $\mathbf{w}, \mathbf{z} \in \mathbb{Z}_\ell^2$ and set $\mathbf{v}_- = \mathbf{w} - \mathbf{z}\sqrt{\varepsilon}$. As g has its coefficients in $\mathbb{Z}_\ell$, the vector $\mathbf{v}_-$ is an eigenvector for g, associated with the eigenvalue $a - b\sqrt{\varepsilon}$. The projections of $\mathbf{v}_\pm$ in $\left(\mathbb{F}_\ell\left[\sqrt{[\varepsilon]}\right]\right)^2$ are therefore nonzero eigenvectors of $[g]$ corresponding to different eigenvalues, hence they are linearly independent. It follows that $\mathbf{w} = \frac{\mathbf{v}_+ + \mathbf{v}_-}{2}, \mathbf{z} = \frac{\mathbf{v}_+ - \mathbf{v}_-}{2\sqrt{\varepsilon}}$ are independent modulo $\ell\mathbb{Z}_\ell[\sqrt{\varepsilon}]$, and since $\mathbf{w}, \mathbf{z}$ lie in $\mathbb{Z}_\ell^2$ they are a fortiori independent modulo ℓ. The matrix $\left(\mathbf{z} \mid \mathbf{w}\right)$ is then invertible modulo ℓ, so it lies in $\mathrm{GL}_2(\mathbb{Z}_\ell)$ and can be used as base-change matrix. It is now straightforward to check that in this basis the element g is represented by the matrix $\begin{pmatrix} a & b\varepsilon \\ b & a \end{pmatrix}$. Finally notice that a and b are units: if $[b] = 0$ or $[a] = 0$ it is easy to check that the order of $G(\ell)$ divides 4, contradicting the assumptions. □

We can also assume that G contains $-\mathrm{Id}$, since replacing G with $G \cdot \{\pm \mathrm{Id}\}$ alters neither the derived subgroup nor the special Lie algebra of G. By lemma 1.4.4 the algebra $L(G)$ admits a basis of the form $\Theta(g), \Theta(g_2), \Theta(g_3)$, where g is as above and g_2, g_3 are in G. We write in coordinates

$$g_2 = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}, \Theta(g_2) = \begin{pmatrix} \frac{y_{11} - y_{22}}{2} & y_{12} \\ y_{21} & -\frac{y_{11} - y_{22}}{2} \end{pmatrix},$$

$$g_3 = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix}, \Theta(g_3) = \begin{pmatrix} \frac{z_{11} - z_{22}}{2} & z_{12} \\ z_{21} & -\frac{z_{11} - z_{22}}{2} \end{pmatrix}.$$

1.4.4.1 Projection operators, φ-stable subalgebras

Recall that φ denotes $x \mapsto g^{-1}xg$. Following our general strategy we now describe projection operators associated with the action of φ and φ-stable subalgebras of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

Lemma 1.4.8. *Let $E, F \in \mathbb{Z}_\ell$. If the matrix $\begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix}$ belongs to $L(N(G))$, then $L(N(G))$ also contains*

$$\begin{pmatrix} -F & 0 \\ 0 & F \end{pmatrix}, \begin{pmatrix} -E & 0 \\ 0 & E \end{pmatrix}, \begin{pmatrix} 0 & -\varepsilon E \\ E & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & -\varepsilon F \\ F & 0 \end{pmatrix}.$$

Proof. We know from lemma 1.4.5 that $L(N(G))$ is φ-stable, so the identity

$$\frac{1}{2ab}\left(\varphi\begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix} - (a^2 + b^2\varepsilon)\begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix}\right) = \begin{pmatrix} -\varepsilon E & -\varepsilon F \\ F & \varepsilon E \end{pmatrix} \tag{1.1}$$

shows that $\begin{pmatrix} -\varepsilon E & -\varepsilon F \\ F & \varepsilon E \end{pmatrix}$ is in $L(N(G))$. At least one of F/E and E/F is an ℓ-adic integer, and we can assume it is F/E (the other case being perfectly analogous). In particular we have $v_\ell(F) \geq v_\ell(E)$. It follows that $L(N(G))$ contains

$$\frac{F}{E}\begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix} - \begin{pmatrix} -\varepsilon E & -\varepsilon F \\ F & \varepsilon E \end{pmatrix} = \begin{pmatrix} \frac{\varepsilon E^2 - F^2}{E} & 0 \\ 0 & -\frac{\varepsilon E^2 - F^2}{E} \end{pmatrix}.$$

If $v_\ell(F) > v_\ell(E)$ we have $v_\ell(\varepsilon E^2 - F^2) = 2v_\ell(E)$, while if $v_\ell(F) = v_\ell(E)$ we can write

$$F = \ell^{v_\ell(E)}\zeta,\ E = \ell^{v_\ell(E)}\gamma,$$

where ζ, γ are not zero modulo ℓ. In this second case we have $\varepsilon E^2 - F^2 = \ell^{2v_\ell(E)}\left(\varepsilon\gamma^2 - \zeta^2\right)$, and $\left(\varepsilon\gamma^2 - \zeta^2\right)$ does not vanish modulo ℓ since $[\varepsilon]$ is not a square in $\mathbb{F}_\ell^\times$. Hence $v_\ell(\varepsilon E^2 - F^2) = 2v_\ell(E)$ holds in any case, and (due to the denominator E) we have found in $L(N(G))$ a matrix whose off-diagonal coefficients vanish and whose diagonal coefficients have the same valuation as E. By the stability of $L(N(G))$ under multiplication by ℓ-adic units we have thus proved that $L(N(G))$ contains $\begin{pmatrix} -E & 0 \\ 0 & E \end{pmatrix}$. Identity (1.1) applied to this element shows that $L(N(G))$ also contains $\begin{pmatrix} 0 & -\varepsilon E \\ E & 0 \end{pmatrix}$, hence by difference $\begin{pmatrix} -F & 0 \\ 0 & F \end{pmatrix}$ is in $L(N(G))$ as well. Applying equation (1.1) to this last matrix we finally deduce that $L(N(G))$ also contains $\begin{pmatrix} 0 & -\varepsilon F \\ F & 0 \end{pmatrix}$. □

Lemma 1.4.9. *Let E, F be elements of $\mathbb{Z}_\ell$ satisfying $\min\{v_\ell(F), v_\ell(E)\} \leq s$. If $\begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix}$ belongs to $L(N(G))$, then $L(N(G))$ contains $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.*

Proof. Suppose $v_\ell(F) \leq s$, the other case being similar. The special Lie algebra $L(N(G))$ contains $\begin{pmatrix} -F & 0 \\ 0 & F \end{pmatrix}$, $\begin{pmatrix} 0 & -\varepsilon F \\ F & 0 \end{pmatrix}$ by the previous lemma, so (given that $v_\ell(F) \leq s$) it also contains $\ell^s\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\ell^s\begin{pmatrix} 0 & -\varepsilon \\ 1 & 0 \end{pmatrix}$. Taking the commutator of these two elements yields another element of

$L(N(G))$, namely

$$\left[\ell^s\begin{pmatrix}0 & -\varepsilon\\ 1 & 0\end{pmatrix}, \ell^s\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right] = \ell^{2s}\begin{pmatrix}0 & 2\varepsilon\\ 2 & 0\end{pmatrix}.$$

Finally, since

$$\frac{1}{2}\ell^{2s}\begin{pmatrix}0 & 2\varepsilon\\ 2 & 0\end{pmatrix} + \ell^{2s}\begin{pmatrix}0 & -\varepsilon\\ 1 & 0\end{pmatrix} = \ell^{2s}\begin{pmatrix}0 & 0\\ 2 & 0\end{pmatrix},$$

it is immediately checked that $L(N(G))$ contains a basis of $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$ as desired. □

1.4.4.2 The case when $g_2, g_3 \notin N(G)$.

Let us assume for now that $g_i \notin N(G)$ and $-g_i \notin N(G)$ for $i = 2, 3$. We will deal later with the case when some of these elements already belong to $N(G)$. Given that by hypothesis $L(G)$ contains $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$ we must have a representation

$$\ell^s\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \sum_{i=1}^{3}\lambda_i\Theta(g_i)$$

for certain scalars $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{Z}_\ell$. However, the diagonal coefficients of $\Theta(g)$ vanish, therefore there exists an index $i \in \{2, 3\}$ such that $v_\ell \circ \pi_{11}(\Theta(g_i)) \leq s$. Renumbering g_2, g_3 if necessary we can assume $i = 2$. In coordinates, the condition $v_\ell \circ \pi_{11}(\Theta(g_2)) \leq s$ becomes $v_\ell(y_{11} - y_{22}) \leq s$.

Now since $[g]$ generates $G(\ell)$ there is an integer k such that $[g]^{-k} = [g_2]$ in $G(\ell)$; in other words, both g_2g^k and g^kg_2 are trivial modulo ℓ and therefore belong to $N(G)$. It is immediate to check that the matrix g^k is of the form $\begin{pmatrix}c & d\varepsilon\\ d & c\end{pmatrix}$ for certain $c, d \in \mathbb{Z}_\ell$. Now if d is 0 modulo ℓ, then (since $c^2 - \varepsilon d^2 \equiv 1 \pmod{\ell}$) we have $c \equiv \pm 1 \pmod{\ell}$, so either g_2 or $-g_2$ reduces to the identity modulo ℓ and is therefore in $N(G)$, contradicting our assumption. Hence d is an ℓ-adic unit. We then introduce

$$g_4 = \begin{pmatrix}c & d\varepsilon\\ d & c\end{pmatrix}\begin{pmatrix}y_{11} & y_{12}\\ y_{21} & y_{22}\end{pmatrix},\ g_5 = \begin{pmatrix}y_{11} & y_{12}\\ y_{21} & y_{22}\end{pmatrix}\begin{pmatrix}c & d\varepsilon\\ d & c\end{pmatrix}.$$

By construction g_4 and g_5 are elements of $N(G)$, whence $\Theta(g_4), \Theta(g_5)$ are elements of $L(N(G))$. In particular $L(N(G))$ contains their difference

$$\Theta(g_4) - \Theta(g_5) = g_4 - g_5 = \begin{pmatrix}-d(y_{12} - \varepsilon y_{21}) & d\varepsilon\,(-y_{11} + y_{22})\\ d\,(y_{11} - y_{22}) & d\,(y_{12} - \varepsilon y_{21})\end{pmatrix},$$

where (given that d, ε are ℓ-adic units) $v_\ell \circ \pi_{21}(\Theta(g_4) - \Theta(g_5)) \leq s$ and $v_\ell \circ \pi_{12}(\Theta(g_4) - \Theta(g_5)) \leq s$. Applying lemma 1.4.9 to the element $\Theta(g_4) - \Theta(g_5)$ we have just constructed we therefore deduce $L(N(G)) \supseteq \ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$ as desired.

1.4.4.3 The case when one generator belongs to $N(G)$.

Let $x = \begin{pmatrix}x_{11} & x_{12}\\ x_{21} & -x_{11}\end{pmatrix}$ denote any element of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. It is easy to check that

$$\frac{1}{2ab}\left((3 + 4\varepsilon b^2)(\varphi x - x) - \varphi(\varphi x - x)\right) = \begin{pmatrix}x_{12} - \varepsilon x_{21} & 2\varepsilon x_{11}\\ -2x_{11} & -x_{12} + \varepsilon x_{21}\end{pmatrix},$$

and furthermore if x belongs to $L(N(G))$, then $\begin{pmatrix} x_{12}-\varepsilon x_{21} & 2\varepsilon x_{11} \\ -2x_{11} & -x_{12}+\varepsilon x_{21} \end{pmatrix}$ is in $L(N(G))$ as well.

Suppose now that either g_2 or $-g_2$ (resp. g_3 or $-g_3$) belongs to $N(G)$. Since $\Theta(-g_i) = -\Theta(g_i)$ we can assume that g_2 (resp. g_3) itself belongs to $N(G)$. Take $\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{pmatrix}$ to be $\Theta(g_2)$ (resp. $\Theta(g_3)$). Subtracting $\frac{x_{21}}{b}\Theta(g_1)$ from $\Theta(g_2)$ we get $\begin{pmatrix} x_{11} & x_{12}-\varepsilon x_{21} \\ 0 & -x_{11} \end{pmatrix} \in L(G)$, and since we know that

$$\Theta(g_2) - \frac{\pi_{21}(\Theta(g_2))}{b}\Theta(g_1),\ \Theta(g_3) - \frac{\pi_{21}(\Theta(g_3))}{b}\Theta(g_1)$$

together span $\ell^s \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \oplus \ell^s \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we see that at least one among the coefficients of the matrix $\Theta(g_2) - \frac{\pi_{21}(\Theta(g_2))}{b}\Theta(g_1) = \Theta(g_2) - \frac{x_{21}}{b}\Theta(g_1)$ must have valuation at most s, that is $\min\{v_\ell(x_{11}), v_\ell(x_{12}-\varepsilon x_{21})\} \leq s$. We now apply lemma 1.4.9 to $\begin{pmatrix} x_{12}-\varepsilon x_{21} & 2\varepsilon x_{11} \\ -2x_{11} & -x_{12}+\varepsilon x_{21} \end{pmatrix}$, which is in $L(N(G))$, to deduce $L(N(G)) \supseteq \ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, and we are done.

1.4.5 Optimality

The following examples show that it is neither possible to extend theorem 1.4.2 to the exceptional case nor to improve the exponent $2s$.

Proposition 1.4.10. *Let ℓ be a prime $\equiv 1 \pmod 4$. For every $t \geq 1$ there exists a closed subgroup G of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ whose special Lie algebra is $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ and whose maximal pro-ℓ subgroup is contained in $\mathcal{B}_\ell(t)$.*

Proof. Notice that the following six elements form a finite subgroup H of $\mathrm{PSL}_2(\mathbb{Z}[i])$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \begin{pmatrix} -i & i \\ 0 & i \end{pmatrix}, \begin{pmatrix} i & 0 \\ i & -i \end{pmatrix},$$

and that H is isomorphic to S_3: indeed, it is the group of permutations of $\{0, 1, \infty\} \subset \mathbb{P}^1(\mathbb{Z}[i])$. The inverse image $\tilde{H}$ of H in $\mathrm{SL}_2(\mathbb{Z}[i])$ is therefore a finite group of cardinality 12. Now since $\ell \equiv 1 \pmod 4$ there is a square root of -1 in $\mathbb{Z}_\ell$, so $\mathbb{Z}[i] \hookrightarrow \mathbb{Z}_\ell$ and $\tilde{H} \hookrightarrow \mathrm{SL}_2(\mathbb{Z}_\ell)$. Consider $G = \tilde{H} \cdot \mathcal{B}_\ell(t) \subset \mathrm{SL}_2(\mathbb{Z}_\ell)$. It is clear that $\mathcal{B}_\ell(t)$ is normal in G. Since $\frac{G}{\mathcal{B}_\ell(t)}$ is isomorphic to a quotient of $\tilde{H}$ (and therefore has order prime to ℓ), the subgroup $\mathcal{B}_\ell(t)$ is clearly the maximal pro-ℓ subgroup of G. Furthermore, the special Lie algebra of G contains the three elements

$$\Theta\left(\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\right) = \begin{pmatrix} -1/2 & 1 \\ -1 & 1/2 \end{pmatrix}, \Theta\left(\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\right) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \Theta\left(\begin{pmatrix} i & 0 \\ i & -i \end{pmatrix}\right) = \begin{pmatrix} i & 0 \\ i & -i \end{pmatrix},$$

that are readily checked to be a basis of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. □

On the other hand, the following example shows that there exist subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ such that $L(G)$ contains $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$, but $L(N(G))$ only contains $\ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Fix $s \geq 1$, an integer $N > 4$ and a prime ℓ congruent to 1 modulo N; then $\mathbb{Z}_\ell^\times$ contains a primitive N-th root of unity a, and we

let $g = \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix}$. The module $M = \ell^s \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \oplus \ell^s \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \oplus \ell^{2s} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is a Lie subalgebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$, so by Theorem 3.4 of [97]

$$H = \left\{ x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; \mathrm{tr}(x) \equiv 2 \pmod{\ell^{2s}}, \Theta(x) \in M \right\}$$

is a pro-ℓ group with special Lie algebra M. Let G be the group generated by g and H. Up to units $\Theta(g)$ is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, so $L(G)$ contains all of $\ell^s \mathfrak{sl}_2(\mathbb{Z}_\ell)$. On the other hand, H is normal in G: one simply needs to check that $g^{-1} M g = M$, and this is obvious from the equality

$$g^{-1} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{pmatrix} g = \begin{pmatrix} x_{11} & \frac{x_{12}}{a^2} \\ a^2 x_{21} & -x_{11} \end{pmatrix}.$$

Finally, H is maximal among the pro-ℓ subgroups of G, since G/H is a quotient of $\langle g \rangle \cong \mathbb{Z}/N\mathbb{Z}$, hence of order prime to ℓ. Therefore $N(G) = H$ and $L(N(G)) = L(H) = M$ contains $\ell^t \mathfrak{sl}_2(\mathbb{Z}_\ell)$ only for $t \geq 2s$.

1.4.6 Proof of theorem 1.4.2

We now prove (1) of theorem 1.4.2 by reducing it to the corresponding statement in theorem 1.4.1. As G and $\mathrm{Sat}(G)$ have the same special Lie algebra and derived subgroup we can assume $G = \mathrm{Sat}(G)$. As G is saturated and satisfies the condition on the determinant, we know from lemma 1.3.17 that $G = \mathrm{Sat}(H)$ for $H = G^{\det=1}$. By the same lemma we also have $L(H) = L(G)$ and $G' = H'$. By assumption H satisfies the hypotheses of theorem 1.4.1 (1), so H has property $(\star)$. As $L(G) = L(H)$ contains $\ell^s \mathfrak{sl}_2(\mathbb{Z}_\ell)$ we deduce that $L_0 = L(N(H))$ contains $\ell^{2s} \mathfrak{sl}_2(\mathbb{Z}_\ell)$, and since $N(H)$ is a pro-ℓ group we can apply theorem 1.3.9 to it. In order to do so we need to estimate $C(N(H)) = \mathrm{tr}\,(L_0 \cdot L_0)$ and $[L_0, L_0]$. Note that

$$C(N(H)) \ni \mathrm{tr}\left(\ell^{2s} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \ell^{2s} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) = 2\ell^{4s},$$

so given that ℓ is odd we have $C(L_0) \supseteq (2\ell^{4s}) = (\ell^{4s})$. Likewise,

$$[L_0, L_0] \supseteq [\ell^{2s} \mathfrak{sl}_2(\mathbb{Z}_\ell), \ell^{2s} \mathfrak{sl}_2(\mathbb{Z}_\ell)] = \ell^{4s} \mathfrak{sl}_2(\mathbb{Z}_\ell),$$

so the derived subgroup of $N(H)$ (which is clearly included in $H' = G'$) is

$$N(H)' = \left\{ x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; \mathrm{tr}\, x - 2 \in C(N(H)), \Theta(x) \in [L_0, L_0] \right\},$$

and by the above it contains

$$\left\{ x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; \mathrm{tr}\, x \equiv 2 \pmod{\ell^{4s}}, \Theta(x) \equiv 0 \pmod{\ell^{4s}} \right\} \supseteq \mathcal{B}_\ell(4s),$$

which concludes the proof of (1).

We are now left with the task of proving (2). Consider first the map

$$G \xrightarrow{\det} \mathbb{Z}_\ell^\times \to \frac{\mathbb{Z}_\ell^\times}{\mathbb{Z}_\ell^{\times 2}} \cong \frac{\mathbb{Z}}{2\mathbb{Z}}$$

and let G_1 be its kernel: then $[G : G_1] \leq 2$, so we can replace G with G_1 and assume that the condition on the determinant is satisfied. We are reduced to showing that, under this hypothesis, either $G' = \mathrm{SL}_2(\mathbb{Z}_\ell)$ or there exists a subgroup H of index at most 12 that satisfies the right

conditions on $\mathrm{Sat}(H)^{\det=1}$. For notational simplicity we let π denote the projection map $G \to G(\ell)$. We now distinguish cases according to ℓ and $G(\ell)$ (cf. theorem 1.3.13):

- if $\ell \geq 5$ and $G(\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$, then it follows from lemma 1.3.15 that $G' = \mathrm{SL}_2(\mathbb{Z}_\ell)$.
- if $\ell = 3$ we let S denote either a 3-Sylow of $G(3)$, if the order of $G(3)$ is a multiple of 3, or the trivial group $\{\mathrm{Id}\}$, if it is not. Notice that $G(3)$ is a subgroup of $\left\{g \in \mathrm{GL}_2(\mathbb{F}_3) \;\middle|\; \det(g) \text{ is a square}\right\}$, which has order 24, so the index $[G(3) : S]$ is at most 8. We set $H = \pi^{-1}(S)$. It is clear that $[G : H] \leq 8$, and H satisfies the conditions in (1) by remark 1.4.3, because $(\mathrm{Sat}\, H)^{\det=1}(3)$ is either $\{\pm \mathrm{Id}\}$ or a group of order 6.
- if $G(\ell)$ is exceptional, then by lemma 1.3.14 there exists a cyclic subgroup B of $\mathbb{P}G(\ell)$ with $[\mathbb{P}G(\ell) : B] \leq 12$: such a B can be taken to have order 3 (resp. 5) if $\mathbb{P}G(\ell)$ is isomorphic to A_4 or S_4 (resp. to A_5). Fix a generator $[b]$ of B and let ξ be the composition $G \to G(\ell) \to \mathbb{P}G(\ell)$. We set $H := \xi^{-1}(B)$; it is clear that $[G : H] \leq 12$. Let now $b \in G(\ell)$ be an element that maps to $[b]$ in B, and let m be the (odd) order of $[b]$. We know that $\det b$ is a square in $\mathbb{F}_\ell^\times$, hence there exists a $\lambda \in \mathbb{F}_\ell^\times$ such that $\det(\lambda b) = 1$. Notice now that $(\lambda b)^m$ is a homothety (it projects to the trivial element in $\mathbb{P}G(\ell)$) and has determinant 1, so it is either Id or $-\mathrm{Id}$; replacing λ by $-\lambda$ if necessary, we can assume that $(\lambda b)^m = -\mathrm{Id}$. By construction, every element in $\left(\mathrm{Sat}(H)^{\det=1}\right)(\ell) = \mathrm{Sat}(H(\ell))^{\det=1}$ can be written as $\pm(\lambda b)^n$ for some $n \in \mathbb{N}$ and for some choice of sign. Now using the fact that $(\lambda b)^m = -\mathrm{Id}$ we see that $\left(\mathrm{Sat}(H)^{\det=1}\right)(\ell)$ is cyclic, generated by λb: since the order of λb is either 6 or 10, H satisfies the conditions in (1) by remark 1.4.3.
- if $G(\ell)$ is contained in a (split or nonsplit) Cartan subgroup then the same is true for the group $\left(\mathrm{Sat}(G)^{\det=1}\right)(\ell)$. If $\left(\mathrm{Sat}(G)^{\det=1}\right)(\ell)$ does not have order 4 we are done, so suppose it does. Then $\mathbb{P}G(\ell)$ has at most 4 elements, and we can take

$$H = \ker\left(G \to G(\ell) \to \mathbb{P}G(\ell)\right) :$$

this H has index at most 4 in G, and $H(\ell)$ has trivial image in $\mathbb{P}\mathrm{GL}_2(\mathbb{F}_\ell)$, so $H(\ell)$ is contained in the homotheties subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$. Therefore $(\mathrm{Sat}(H))^{\det=1}(\ell) = \mathrm{Sat}(H(\ell))^{\det=1} = \{\pm \mathrm{Id}\}$ and H satisfies the conditions in (1).
- if $G(\ell)$ is contained in the normalizer of a (split or nonsplit) Cartan subgroup $\mathcal{C}$, but not in $\mathcal{C}$ itself, then G has a subgroup G_1 of index 2 whose image modulo ℓ is contained in $\mathcal{C}$, and we are reduced to the Cartan case.
- if $G(\ell)$ is contained in a Borel subgroup, then the same is true for $\mathrm{Sat}(G)^{\det=1}(\ell)$. To ease the notation we set $G_2 = \mathrm{Sat}(G)^{\det=1}$. We can also assume that ℓ divides the order of $G(\ell)$ (hence that of $G_2(\ell)$ as well), for otherwise we are back to the (split) Cartan case. Now if $|G_2/N(G_2)| \neq 4$ we can set $H = G$; if, on the contrary, $|G_2/N(G_2)| = 4$ we consider the group morphism

$$\begin{array}{rccccc} \tau : & G & \to & G(\ell) & \to & \mathbb{F}_\ell^\times \\ & g & \mapsto & [g] = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} & \mapsto & a/c. \end{array}$$

Every $g \in G$ is of the form λg_2 for suitable $\lambda \in \mathbb{Z}_\ell^\times$ and $g_2 \in G_2$, and since $\tau(\lambda g_2) = \tau(g_2)$ we deduce $\tau(G) = \tau(G_2)$. On the other hand, when restricted to G_2 the function τ becomes

$$g \mapsto [g] = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} \mapsto a^2,$$

and as we have already remarked $g \mapsto [g] = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} \mapsto a$ is the quotient map $G_2 \twoheadrightarrow G_2/N(G_2)$. Hence τ factors through the quotient $G_2/N(G_2)$ and we have $|\tau(G)| = |\tau(G_2)| \mid 4$. We take H to be the kernel of τ. Then it is clear that $[G : H]$ divides 4, and we claim that H satisfies the conditions in (1). To check this last claim, notice first that $H(\ell)$ is a subgroup of $G(\ell)$, so it is contained in a Borel subgroup. We also have $\ker \pi \subseteq H$, so $G/H \cong \frac{G/\ker \pi}{H/\ker \pi} = \frac{G(\ell)}{H(\ell)}$; in particular $[G(\ell) : H(\ell)]$ divides 4, and therefore the order of $H(\ell)$ is divisible by ℓ. Finally, any matrix $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ in $H(\ell)$ satisfies $a/c = 1$ by construction, so the intersection $\mathrm{Sat}(H(\ell)) \cap \mathrm{SL}_2(\mathbb{F}_\ell)$ consists of matrices $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ with $a = c$ and $ac = 1$, so $a = c = \pm 1$. This implies that the quotient of $\mathrm{Sat}(H)^{\det=1}(\ell)$ by its ℓ-Sylow has at most 2 elements, and since this quotient is exactly $\mathrm{Sat}(H)^{\det=1}/N\left(\mathrm{Sat}(H)^{\det=1}\right)$ the result follows. □

Remark 1.4.11. For future applications, we remark that the same proof shows that the inequality $[G : H] \leq 24$ appearing in theorem 1.4.2 (2) can be replaced by the condition $[G : H] \mid 48$, and even by $[G : H] \mid 24$ if in addition G satisfies $\det(G) \subseteq \mathbb{Z}_\ell^{\times 2}$.

1.5 Recovering G from $L(G)$, when $\ell = 2$

We now consider closed subgroups of $\mathrm{GL}_2(\mathbb{Z}_2)$, and endeavour to show results akin to those of the previous section. For $\mathrm{GL}_2(\mathbb{Z}_2)$ the statement is as follows:

Theorem 1.5.1. *Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_2)$.*

1. *Suppose that $G(4)$ is trivial and $\det(G) \equiv 1 \pmod 8$. The following implication holds for all positive integers n: if $L(G)$ contains $2^n\mathfrak{sl}_2(\mathbb{Z}_2)$, then the derived subgroup G' of G contains the principal congruence subgroup $\mathcal{B}_2(12n + 2)$.*

2. *Without any assumption on G, the subgroup*
$$H = \ker(G \to G(4)) \cap \ker\left(G \to G(8) \overset{\det}{\to} (\mathbb{Z}/8\mathbb{Z})^\times\right)$$
satisfies $[G : H] \leq 2 \cdot 96 = 192$ and the conditions in (1).

Note that (2) is immediate: the order of $\mathrm{GL}_2(\mathbb{Z}/4\mathbb{Z})$ is 96, and once we demand that $G(4)$ is trivial the determinant modulo 8 can only take two different values. As in the previous section, the core of the problem lies in understanding the subgroups of $\mathrm{SL}_2(\mathbb{Z}_2)$, so until the very last paragraph of this section the letter G will denote a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$. In view of the result we want to prove, we will also enforce the assumption that G has trivial reduction modulo 4; indeed in this context the relevant statement is:

Theorem 1.5.2. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$ whose reduction modulo 4 is trivial, and let s be an integer no less than 2. If $L(G)$ contains $2^s\mathfrak{sl}_2(\mathbb{Z}_2)$, then G contains $\mathcal{B}_2(6s)$.*

The idea of the proof is quite simple: despite the fact there is in general no reason why $\Theta(G)$ should be a group under addition, we will show that for every pair x, y of elements of $\Theta(G)$ it is possible

to find an element that is reasonably close to $x+y$ and that lies again in $\Theta(G)$. The error term will turn out to be quadratic in x and y, which is not quite good enough by itself, since a correction of this order of magnitude could still be large enough to destroy any useful information about $x+y$; the technical step needed to make the argument work is that of multiplying all the elements we have to deal with by a power of 2 large enough that the quadratic error term becomes negligible with respect to the linear part. The rest of the proof is really just careful bookkeeping of the correction terms appearing in the various addition formulas. We shall continue using the notation from the previous section:

Notation. For $x \in L := L(G)$ we set $\pi_{ij}(x) = x_{ij}$, the coefficient in the i-th row and j-th column of the matrix representation of x in $\mathfrak{sl}_2(\mathbb{Z}_2)$. The maps π_{ij} are linear and continuous.

We start with a compactness lemma. Our arguments only yield (arbitrarily good) approximations of elements of $\Theta(G)$, and we need to know that this is enough to show that the matrices we are approximating actually belong to $\Theta(G)$.

Lemma 1.5.3. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$, g be an element of G, and $e \geq 2$. Suppose that $\Theta(g) \equiv 0 \pmod{2^e}$: then $\mathrm{tr}(g) - 2$ is divisible by 2^{2e}. Moreover $\Theta^{-1} : \Theta(G) \cap 2^2\mathfrak{sl}_2(\mathbb{Z}_2) \to G$ is well defined and continuous, and the intersection $\Theta(G) \cap 2^2\mathfrak{sl}_2(\mathbb{Z}_2)$ is compact.*

Proof. Write $\Theta(g) = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ and $g = \dfrac{\mathrm{tr}(g)}{2}\,\mathrm{Id} + \Theta(g)$. As G is a subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$, we have the identity

$$1 = \det g = \det\left(\frac{\mathrm{tr}(g)}{2}\,\mathrm{Id} + \Theta(g)\right) = \left(\frac{\mathrm{tr}(g)}{2}\right)^2 - a^2 - bc.$$

Furthermore G (hence g) is trivial modulo 4 by assumption, so an immediate calculation shows that $1 = \det(g) \equiv 1 + (\mathrm{tr}(g) - 2) \pmod 8$. It follows that $\frac{\mathrm{tr}(g)}{2}$ is the unique solution to the equation $\lambda^2 = 1 + a^2 + bc$ that is congruent to 1 modulo 4, hence $\dfrac{\mathrm{tr}(g)}{2} = \sqrt{1 + a^2 + bc} = \displaystyle\sum_{j=0}^{\infty} \binom{1/2}{j} (a^2 + bc)^j$ by lemma 1.3.2. Given that $a^2 + bc \equiv 0 \pmod{2^{2e}}$ and $2e > 3$, using again lemma 1.3.2 we find

$$v_2\left(\mathrm{tr}(g) - 2\right) = v_2\left(2\left(\frac{\mathrm{tr}(g)}{2} - 1\right)\right) = 1 + v_2\left(\sqrt{1 + (a^2 + bc)} - 1\right) \geq 2e.$$

The case $e = 2$ of the above computation shows that every $x \in 2^2\mathfrak{sl}_2(\mathbb{Z}_2)$ admits exactly one inverse image in $\mathrm{SL}_2(\mathbb{Z}_2)$ that reduces to the identity modulo 4, so $\Theta : \mathcal{B}_2(2) \to 2^2\mathfrak{sl}_2(\mathbb{Z}_2)$ is a continuous bijection: we have just described the (two-sided) inverse, so we only need to check that the image of $\mathcal{B}_2(2)$ under Θ does indeed land in $2^2\mathfrak{sl}_2(\mathbb{Z}_2)$. We have to show that if $g = \begin{pmatrix} d & b \\ c & e \end{pmatrix}$ is any element of $\mathcal{B}_2(2)$, then $\Theta(g) = \begin{pmatrix} \frac{d-e}{2} & b \\ c & \frac{e-d}{2} \end{pmatrix}$ has all its coefficients divisible by 4. This is obvious for b and c. For the diagonal ones, note that $de - bc = 1$, so $de \equiv 1 \pmod 8$ and hence $d \equiv e \pmod 8$ and $\frac{d-e}{2} \equiv 0 \pmod 4$ as required. Observe now that $a^2 + bc = \frac{1}{2}\,\mathrm{tr}\left(\Theta(g)^2\right)$, so we can write

$$\Theta^{-1}(x) = x + \sqrt{1 + \frac{1}{2}\,\mathrm{tr}(x^2)} \cdot \mathrm{Id},$$

which is manifestly continuous. Therefore Θ establishes a homeomorphism between $\mathcal{B}_2(2)$ and $2^2\mathfrak{sl}_2(\mathbb{Z}_2)$.

In particular, we have a well-defined and continuous map $\Theta^{-1} : \Theta(G) \cap 2^2\mathfrak{sl}_2(\mathbb{Z}_2) \to G$, and we finally deduce that the intersection $\Theta(G) \cap 2^2\mathfrak{sl}_2(\mathbb{Z}_2) = \Theta(G \cap \mathcal{B}_2(2))$ is compact, since this is true for $G \cap \mathcal{B}_2(2)$ and Θ is continuous. □

The core of the proof of theorem 1.5.2 is contained in the following lemma:

Lemma 1.5.4. *Let e_1, e_2 be integers not less than 2 and x_1, x_2 be elements of $\Theta(G)$. Suppose that $x_1 \equiv 0 \pmod{2^{e_1}}$ and $x_2 \equiv 0 \pmod{2^{e_2}}$: then $\Theta(G)$ contains an element y congruent to $x_1 + x_2$ modulo $2^{e_1+e_2-1}$. If, furthermore, both x_1 and x_2 are in upper-triangular form, then we can find such a y having the same property.*

Proof. Write $x_1 = \Theta(g_1)$, $x_2 = \Theta(g_2)$ and set $y = \Theta(g_1 g_2)$. Applying lemma 1.3.10 we find

$$2\,(y - x_1 - x_2) = [x_1, x_2] + (\mathrm{tr}(g_1) - 2)x_2 + (\mathrm{tr}(g_2) - 2)x_1.$$

Consider the 2-adic valuation of the various terms on the right. The commutator $[x_1, x_2]$ is clearly 0 modulo $2^{e_1+e_2}$. We also have $\mathrm{tr}(g_1) - 2 \equiv 0 \pmod{2^{2e_1}}$ and $\mathrm{tr}(g_2) - 2 \equiv 0 \pmod{2^{2e_2}}$ by lemma 1.5.3, so the last two terms are divisible respectively by $2^{2e_1+e_2}$ and $2^{e_1+2e_2}$. It follows that the right hand side of this equality is zero modulo $2^{e_1+e_2}$, and dividing by 2 we get the first statement in the lemma.
For the last claim simply note that if x_1, x_2 are upper-triangular then the same is true for all of the error terms, so $y = x_1 + x_2 +$ (triangular error terms) is indeed triangular. □

As a first application, we show that the image of Θ is stable under multiplication by 2 (up to units):

Lemma 1.5.5. *Let $x \in \Theta(G)$ and $m \in \mathbb{N}$. There exists a unit $\lambda \in \mathbb{Z}_2^\times$ such that $\lambda \cdot 2^m x$ again belongs to $\Theta(G)$.*

Proof. Clearly there is nothing to prove for $m = 0$, so let us start with the case $m = 1$. Write $x = \Theta(g)$ for a certain $g \in G$. By our assumptions on G, the trace of g is congruent to 2 modulo 4, so $\lambda = \frac{\mathrm{tr}(g)}{2}$ is a unit in $\mathbb{Z}_2$. We can therefore form $\tilde{g} = \frac{1}{\lambda} g$, which certainly exists as a matrix in $\mathrm{GL}_2(\mathbb{Z}_2)$, even though it does not necessarily belong to G. Our choice of $\tilde{g}$ is made so as to ensure $\mathrm{tr}(\tilde{g}) = 2$, so the formula given in lemma 1.3.10 (applied with $g_1 = g_2 = \tilde{g}$) yields

$$2\left(\Theta\left(\tilde{g}^2\right) - \Theta(\tilde{g}) - \Theta(\tilde{g})\right) = [\Theta(\tilde{g}), \Theta(\tilde{g})] + (\mathrm{tr}(\tilde{g}) - 2)\,\Theta(\tilde{g}) + (\mathrm{tr}(\tilde{g}) - 2)\,\Theta(\tilde{g}),$$

where the right hand side vanishes. We deduce $\Theta(\tilde{g}^2) = 2\Theta(\tilde{g})$, and it is now immediate to check that $\Theta(g^2) = \lambda \cdot 2\Theta(g)$, whence the claim for $m = 1$. An immediate induction then proves the general case. □

We now take the first step towards understanding the structure of $\Theta(G)$, namely showing that a suitable basis of L can be found inside $\Theta(G)$. Note that L, being open, is automatically of rank 3.

Lemma 1.5.6. *There exist a basis $\{x_1, x_2, x_3\} \subseteq \Theta(G)$ of L and scalars $\tilde{\sigma}_{21}$, $\tilde{\sigma}_{31}$, $\tilde{\sigma}_{32} \in \mathbb{Z}_2$ with the following properties: $\pi_{21}(x_2 - \tilde{\sigma}_{21}x_1) = 0$, $\pi_{21}(x_3 - \tilde{\sigma}_{31}x_1) = 0$ and*

$$\pi_{21}(x_3 - \tilde{\sigma}_{31}x_1 - \tilde{\sigma}_{32}(x_2 - \tilde{\sigma}_{21}x_1)) = \pi_{11}(x_3 - \tilde{\sigma}_{31}x_1 - \tilde{\sigma}_{32}(x_2 - \tilde{\sigma}_{21}x_1)) = 0.$$

Remark 1.5.7. The slightly awkward equations appearing in the statement of this lemma actually have a simple interpretation: they mean it is possible to subtract a suitable multiple of x_1 from x_2 and x_3 so as to make them upper-triangular, and that it is then further possible to subtract one of the matrices thus obtained from the other so as to leave it with only one nonzero coefficient (in the top right corner).

Proof. This is immediate from lemma 1.3.11, which can be applied identifying $\mathfrak{sl}_2(\mathbb{Z}_2) \cong \mathbb{Z}_2^3$ via $\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \mapsto (c, a, b)$. Note that with this identification the three canonical projections $\mathbb{Z}_2^3 \to \mathbb{Z}_2$ become π_{21}, π_{11} and π_{12} respectively, and the vanishing conditions in the statement become exactly those of lemma 1.3.11. □

As previously mentioned, in order to make the quadratic error terms appearing in lemma 1.5.4 negligible we need to work with matrices that are highly divisible by 2:

Lemma 1.5.8. *Let x_1, x_2, x_3 be a basis of L. There exist elements $y_1, y_2, y_3 \in \Theta(G)$ and units $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{Z}_2^\times$ such that $y_i = \lambda_i \cdot 2^{4s} x_i$ for $i = 1, 2, 3$; in particular y_1, y_2, y_3 are zero modulo 2^{4s}, and the module generated by y_1, y_2, y_3 over $\mathbb{Z}_2$ contains $2^{5s}\mathfrak{sl}_2(\mathbb{Z}_2)$.*

Proof. Everything is obvious (by lemma 1.5.5) except perhaps the last statement. Note that y_1, y_2, y_3 differ from $2^{4s}x_1, 2^{4s}x_2, 2^{4s}x_3$ only by multiplication by units, so these two sets generate over $\mathbb{Z}_2$ the same module M. But the x_i generate $L \supseteq 2^s\mathfrak{sl}_2(\mathbb{Z}_2)$, hence $M = 2^{4s}L$ contains $2^{5s}\mathfrak{sl}_2(\mathbb{Z}_2)$. □

Notation. Let x_1, x_2, x_3 be a basis of L as in lemma 1.5.6, and let y_1, y_2, y_3 be the elements given by lemma 1.5.8 when applied to x_1, x_2, x_3. The properties of the x_i become corresponding properties of the y_i:

- There is a scalar $\sigma_{21} \in \mathbb{Z}_2$ such that
$$y_2 - \sigma_{21} \cdot y_1 = \begin{pmatrix} b_{11} & b_{12} \\ 0 & -b_{11} \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{Z}_2);$$
- there are scalars σ_{31}, σ_{32} such that
$$y_3 - \sigma_{31} y_1 = \begin{pmatrix} d_{11} & d_{12} \\ 0 & -d_{11} \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{Z}_2),$$
$$y_3 - \sigma_{31} y_1 - \sigma_{32}(y_2 - \sigma_{21} \cdot y_1) = \begin{pmatrix} 0 & c_{12} \\ 0 & 0 \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{Z}_2).$$

To ease the notation a little we set
$$t_1 = y_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix}, t_2 = \begin{pmatrix} b_{11} & b_{12} \\ 0 & -b_{11} \end{pmatrix} \text{ and } t_3 = \begin{pmatrix} 0 & c_{12} \\ 0 & 0 \end{pmatrix}.$$
It is clear that $\{t_1, t_2, t_3\}$ and $\{y_1, y_2, y_3\}$ generate the same module M over $\mathbb{Z}_2$, so in particular M contains $2^{5s}\mathfrak{sl}_2(\mathbb{Z}_2)$.

Lemma 1.5.9. *The 2-adic valuations of a_{21}, b_{11} and c_{12} do not exceed $5s$.*

Proof. We can express $\begin{pmatrix} 0 & 0 \\ 2^{5s} & 0 \end{pmatrix}$ as a $\mathbb{Z}_2$-linear combination of t_1, t_2, t_3,

$$\begin{pmatrix} 0 & 0 \\ 2^{5s} & 0 \end{pmatrix} = \lambda_1 t_1 + \lambda_2 t_2 + \lambda_3 t_3,$$

for a suitable choice of $\lambda_1, \lambda_2, \lambda_3$ in $\mathbb{Z}_2$. Comparing the bottom-left coefficients we find $\lambda_1 a_{21} = 2^{5s}$, so $v_2(a_{21}) \leq 5s$ as claimed.

The same argument, applied to the representation of $\begin{pmatrix} 2^{5s} & 0 \\ 0 & -2^{5s} \end{pmatrix}$ (resp. $\begin{pmatrix} 0 & 2^{5s} \\ 0 & 0 \end{pmatrix}$) as a combination of t_1, t_2, t_3, gives $b_{11} | 2^{5s}$ (resp. $c_{12} | 2^{5s}$) and finishes the proof of the lemma. □

For future reference, and since it is easy to lose track of all the notation, we record here two facts we will need later:

Remark 1.5.10. We have $\sigma_{32} = \dfrac{d_{11}}{b_{11}}$ and $v_2(d_{12} - \sigma_{32} b_{12}) = v_2(c_{12}) \leq 5s$.

We now further our investigation of the approximate additive structure of $\Theta(G)$. Since essentially all of the arguments are based on sequences of approximations the following notation will turn out to be very useful.

Notation. We write $a = b + O(2^n)$ if $a \equiv b \pmod{2^n}$.

Lemma 1.5.11. *Let $a_1, a_2 \in \Theta(G) \cap 2^{4s}\mathfrak{sl}_2(\mathbb{Z}_2)$ and $\xi \in \mathbb{Z}_2$. Then $\Theta(G)$ contains an element z congruent to $a_1 - \xi a_2$ modulo 2^{8s-1}. If moreover a_1, a_2 are upper triangular then z can be chosen to have the same property.*

Proof. We construct a sequence $(z_n)_{n \geq 0}$ of elements of $\Theta(G)$ and a sequence $(\xi_n)_{n \geq 0}$ of elements of $\mathbb{Z}_2$ satisfying $\xi_n = \xi + O(2^n)$ and

$$z_n = a_1 - \xi_n a_2 + O\left(2^{8s-1}\right).$$

We can take $z_0 = a_1$ and $\xi_0 = 0$. Given z_n, ξ_n we proceed as follows. If we let $w_n = v_2(\xi_n - \xi)$, then $w_n \geq n$ by the induction hypothesis, and by lemma 1.5.5 we can find a unit λ_n such that $2^{w_n}\lambda_n a_2$ also belongs to $\Theta(G)$. Note that both z_n and $2^{w_n}\lambda_n a_2$ are zero modulo 2^{4s}. Apply lemma 1.5.4 to $(x_1, x_2) = (z_n, 2^{w_n}\lambda_n a_2)$: it yields the existence of an element z_{n+1} of $\Theta(G)$ of the form $z_n + 2^{w_n}\lambda_n a_2 + O\left(2^{8s-1}\right)$. We take $\xi_{n+1} = (\xi_n - 2^{w_n}\lambda_n)$; let us check that ξ_{n+1}, z_{n+1} have the right properties. Clearly

$$z_{n+1} = z_n + 2^{w_n}\lambda_n a_2 + O\left(2^{8s-1}\right) = a_1 - (\xi_n - 2^{w_n}\lambda_n)a_2 + O\left(2^{8s-1}\right).$$

On the other hand the definition of w_n implies that $\xi_n - \xi = 2^{w_n} \cdot \mu_n$ where μ_n is a unit, so

$$\begin{aligned} v_2\left(\xi_{n+1} - \xi\right) &= v_2\left(\left(\xi_n - 2^{w_n}\lambda_n\right) - \xi\right) \\ &= v_2(2^{w_n} \cdot \mu_n - 2^{w_n} \cdot \lambda_n) \\ &= w_n + v_2(\mu_n - \lambda_n) \geq w_n + 1 \geq n + 1, \end{aligned}$$

since μ_n, λ_n are both units and therefore odd. To conclude the proof it is simply enough to take $z = z_{8s-1}$: indeed

$$\begin{aligned} a_1 - \xi a_2 - z_{8s-1} &= a_1 - \xi a_2 - \left(a_1 - \xi_{8s-1} a_2 + O\left(2^{8s-1}\right)\right) \\ &= (\xi_{8s-1} - \xi) a_2 + O\left(2^{8s-1}\right) \\ &= O\left(2^{8s-1}\right) \end{aligned}$$

as required. The proof in the upper-triangular case goes through completely unchanged, simply using the corresponding second part of lemma 1.5.4. □

The above lemma is still not sufficient, since it cannot guarantee that we will ever find a matrix with a coefficient that vanishes exactly. This last remaining obstacle is overcome through the following result:

Lemma 1.5.12. *Let $a_1, a_2 \in \Theta(G) \cap 2^{4s}\mathfrak{sl}_2(\mathbb{Z}_2)$ and $\xi \in \mathbb{Z}_2$. Suppose that for a certain pair (i,j) the (i,j)-th coefficient of $a_1 - \xi a_2$ vanishes while $v_2 \circ \pi_{ij}(a_2) \leq 5s$: then $\Theta(G)$ contains an element z whose (i,j)-th coefficient is zero and that is congruent to $a_1 - \xi a_2$ modulo 2^{7s-1}. If, furthermore, a_1, a_2 are upper-triangular, then this z can be chosen to be upper-triangular as well (while still satisfying $\pi_{ij}(z) = 0$).*

Proof. Let z_0 be the element whose existence is guaranteed by lemma 1.5.11 when applied to a_1, a_2, ξ. We propose to build a sequence $(z_n)_{n\geq 0}$ of elements of $\Theta(G)$ satisfying the following conditions:

1. $z_{n+1} \equiv z_n \pmod{2^{7s-1}}$, and therefore $z_n \equiv z_0 \equiv 0 \pmod{2^{4s}}$;
2. the sequence $w_n = v_2 \circ \pi_{ij}(z_n)$ is monotonically strictly increasing; in particular we have $w_n \geq w_0 \geq 8s - 1$.

Suppose we have constructed z_n, w_n and let $k = v_2 \circ \pi_{ij}(a_2) \leq 5s$. By lemma 1.5.5 we can find a unit λ such that $2^{w_n - k}\lambda a_2$ also belongs to $\Theta(G)$ (note that $w_n \geq 8s - 1 \geq 5s \geq k$). We know that $z_n \equiv 0 \pmod{2^{4s}}$ and $2^{w_n-k}\lambda a_2 \equiv 0 \pmod{2^{w_n-k+4s}}$ (note that $a_2 \equiv 0 \pmod{2^{4s}}$). Apply lemma 1.5.4 to $(x_1, x_2) = (z_n, 2^{w_n-k}\lambda a_2)$: it yields the existence of an element z_{n+1} of $\Theta(G)$ that is congruent to $z_n + 2^{w_n-k}\lambda a_2$ modulo $2^{(4s+w_n-k)+4s-1}$.
We can write $\pi_{ij}(z_n) = 2^{w_n}\mu_n$ and $\pi_{ij}(a_2) = 2^k \xi$ with $\mu_n, \xi \in \mathbb{Z}_2^\times$, so

$$v_2 \circ \pi_{ij}(z_n + 2^{w_n-k}\lambda a_2) = v_2(2^{w_n}\mu_n + 2^{w_n-k}2^k \cdot \xi\lambda) = w_n + v_2(\mu_n + \xi\lambda),$$

and since μ_n, ξ and λ are all odd the last term is at least $w_n + 1$. As k is at most $5s$ by hypothesis we deduce

$$\begin{aligned} w_{n+1} &= v_2 \circ \pi_{ij}(z_{n+1}) \\ &= v_2 \circ \pi_{ij}\left(z_n + 2^{w_n-k}\lambda a_2 + O\left(2^{(4s+w_n-k)+4s-1}\right)\right) \\ &\geq \min\left\{v_2 \circ \pi_{ij}\left(z_n + 2^{w_n-k}\lambda a_2\right), 8s - 1 + w_n - k\right\} \\ &> w_n. \end{aligned}$$

As $2^{w_n-k}\lambda a_2 \equiv 0 \pmod{2^{w_n-k+4s}}$, the difference $z_{n+1} - z_n$ is zero modulo 2^{w_n-s}, hence a fortiori modulo 2^{7s-1} since $w_n \geq w_0 \geq 8s - 1$.

Lemma 1.5.3 says that $\Theta(G) \cap 2^2\mathfrak{sl}_2(\mathbb{Z}_2)$ is compact, so z_n admits a subsequence converging to a certain $z \in \Theta(G)$. By continuity of π_{ij} it is immediate to check that $\pi_{ij}(z) = 0$, and since every z_n is congruent modulo 2^{7s-1} to z_0 the same is true for z. Given that z_0 is congruent to $a_1 - \xi a_2$ modulo 2^{8s-1}, the last assertion follows.
Finally, the upper-triangular case is immediate, since it is clear from the construction that if a_1, a_2 are upper-triangular then the same is true for all the approximations z_n. □

The result we were really aiming for follows at once:

Proposition 1.5.13. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$ whose reduction modulo 2 is trivial, and let s be an integer no less than 2. If $L(G)$ contains $2^s\mathfrak{sl}_2(\mathbb{Z}_2)$, then $\Theta(G)$ contains both an element of the form $\begin{pmatrix} 0 & \tilde{c}_{12} \\ 0 & 0 \end{pmatrix}$, where $v_2(\tilde{c}_{12}) \leq 5s$, and one of the form $\begin{pmatrix} f_{11} & 0 \\ 0 & -f_{11} \end{pmatrix}$, where $v_2(f_{11}) \leq 6s$.*

Proof. We apply lemma 1.5.12 to $a_1 = y_2$, $a_2 = y_1$, $\xi = \sigma_{21}$, $(i,j) = (2,1)$; the hypotheses are satisfied since $y_1 \equiv y_2 \equiv 0 \pmod{2^{4s}}$ and $v_2 \circ \pi_{21}(y_1) \leq 5s$ by lemma 1.5.9. It follows that $\Theta(G)$ contains a matrix $\tilde{b}$ of the form $\begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ 0 & -\tilde{b}_{11} \end{pmatrix}$, where we have $\tilde{b}_{ij} \equiv b_{ij} \pmod{2^{7s-1}}$ for every $1 \leq i,j \leq 2$; in particular, $v_2(\tilde{b}_{11}) \leq 5s$.
The same lemma, applied to $a_1 = y_3, a_2 = y_1$ and $\xi = \sigma_{31}$, implies that $\Theta(G)$ contains a matrix $\tilde{d}$ of the form $\begin{pmatrix} \tilde{d}_{11} & \tilde{d}_{12} \\ 0 & -\tilde{d}_{11} \end{pmatrix}$, where for every i,j we have $\tilde{d}_{ij} \equiv d_{ij} \pmod{2^{7s-1}}$; in particular,

$$v_2(\tilde{d}_{11}) \geq \min\{7s-1, v_2(d_{11})\} \geq v_2(b_{11}) = v_2(\tilde{b}_{11}).$$

Now since $v_2(\tilde{d}_{11}) \geq v_2(\tilde{b}_{11})$ we can find a scalar ζ such that

$$\tilde{d} - \zeta\tilde{b} = \begin{pmatrix} \tilde{d}_{11} & \tilde{d}_{12} \\ 0 & -\tilde{d}_{11} \end{pmatrix} - \zeta \begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ 0 & -\tilde{b}_{11} \end{pmatrix} = \begin{pmatrix} 0 & e_{12} \\ 0 & 0 \end{pmatrix},$$

so applying once again lemma 1.5.12 (more precisely, the version for triangular matrices) we find that $\Theta(G)$ contains a certain matrix $\tilde{e} = \begin{pmatrix} 0 & \tilde{e}_{12} \\ 0 & 0 \end{pmatrix}$, where $\tilde{e}_{12} \equiv e_{12} \pmod{2^{7s-1}}$. Observe now that

$$\zeta = \frac{\tilde{d}_{11}}{\tilde{b}_{11}} = \frac{d_{11} + O\left(2^{7s-1}\right)}{b_{11} + O\left(2^{7s-1}\right)} = \frac{d_{11}}{b_{11}} + O\left(2^{7s-1-v_2(b_{11})}\right) = \frac{d_{11}}{b_{11}} + O\left(2^{2s-1}\right),$$

so upon multiplying by $\tilde{b}_{12}$, which is divisible by 2^{4s}, we obtain the congruence $\zeta\tilde{b}_{12} \equiv \frac{d_{11}}{b_{11}}\tilde{b}_{12} \pmod{2^{6s-1}}$. Since furthermore $\tilde{b}_{12} \equiv b_{12} \pmod{2^{6s-1}}$ we deduce $\zeta\tilde{b}_{12} \equiv \frac{d_{11}}{b_{11}}b_{12} \pmod{2^{6s-1}}$. But then the inequality $v_2(c_{12}) \leq 5s$ (cf. remark 1.5.10) implies

$$\begin{aligned} v_2(\tilde{e}_{12}) &= v_2\left(e_{12} + O\left(2^{7s-1}\right)\right) \\ &= v_2\left(\tilde{d}_{12} - \zeta\tilde{b}_{12} + O\left(2^{7s-1}\right)\right) \\ &= v_2\left(d_{12} - \frac{d_{11}}{b_{11}}b_{12} + O\left(2^{6s-1}\right)\right) \\ &= v_2\left(c_{12} + O\left(2^{6s-1}\right)\right) \\ &\leq 5s. \end{aligned}$$

The existence of the diagonal element is now almost immediate: indeed, we can apply once more lemma 1.5.12 to the difference

$$2^s \begin{pmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ 0 & -\tilde{b}_{11} \end{pmatrix} - \frac{2^s \tilde{b}_{12}}{\tilde{e}_{12}} \begin{pmatrix} 0 & \tilde{e}_{12} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \tilde{b}_{11} & 0 \\ 0 & -\tilde{b}_{11} \end{pmatrix},$$

the hypotheses being satisfied since clearly $2^s \tilde{b} \equiv 0 \pmod{2^{5s}}$ and $v_2(\tilde{e}_{12}) \leq 5s$ for what we have just seen. It follows that $\Theta(G)$ contains a matrix $\begin{pmatrix} f_{11} & 0 \\ 0 & -f_{11} \end{pmatrix}$ congruent to $2^s \begin{pmatrix} \tilde{b}_{11} & 0 \\ 0 & -\tilde{b}_{11} \end{pmatrix}$ modulo 2^{7s-1}, and this is enough to deduce

$$v_2(f_{11}) = v_2(2^s b_{11} + O\left(2^{7s-1}\right)) = s + v_2(b_{11}) \leq 6s.$$

□

We are now ready for the proof of theorem 1.5.2:

Proof of theorem 1.5.2. With all the preliminaries in place this is now quite easy: by proposition 1.5.13 we know that $\Theta(G)$ contains an element of the form $\begin{pmatrix} 0 & \tilde{c}_{12} \\ 0 & 0 \end{pmatrix}$, where $v_2(\tilde{c}_{12}) \leq 5s$, and by the explicit description of Θ^{-1} (lemma 1.5.3) this element must come from $R_{\tilde{c}_{12}} = \begin{pmatrix} 1 & \tilde{c}_{12} \\ 0 & 1 \end{pmatrix} \in G$. Similarly, if we let f denote the diagonal element $\begin{pmatrix} f_{11} & 0 \\ 0 & -f_{11} \end{pmatrix}$, then

$$\Theta^{-1}(f) = \begin{pmatrix} f_{11} & 0 \\ 0 & -f_{11} \end{pmatrix} + \sqrt{1 + \frac{1}{2}\operatorname{tr}(f^2)} \cdot \mathrm{Id}$$

is an operator of the form $D_c = \begin{pmatrix} 1+c & 0 \\ 0 & \frac{1}{c+1} \end{pmatrix}$, where

$$\begin{aligned} v_2(c) &= v_2\left(f_{11} + \sqrt{1 + \frac{1}{2}\operatorname{tr}(f^2)} - 1\right) \\ &= v_2\left(f_{11} + O\left(2^{2v_2(f_{11})-1}\right)\right) \\ &= v_2(f_{11}) \leq 6s. \end{aligned}$$

Observe now that replacing G with G^t, the group $\left\{g^t \mid g \in G\right\}$ endowed with the obvious product $g_1^t \cdot g_2^t = (g_2 g_1)^t$, simply exchanges $L(G)$ for $L(G)^t$, so if $L(G)$ contains the (symmetric) set $2^s \mathfrak{sl}_2(\mathbb{Z}_2)$, then the same is true for $L(G^t)$. Thus G^t contains $R_{2^{5s}}$ and G contains $L_{2^{5s}}$. We have just shown that G contains L_a, R_b and D_c for certain a, b, c of valuation at most $6s$, so it follows from lemma 1.3.4 that G contains $\mathcal{B}_2(6s)$. □

Remark 1.5.14. The above result should be thought of as an analogue of theorem 1.3.9 for $\ell = 2$, even though the present result is actually much weaker. It would of course be interesting to have a complete classification result for pro-2 groups purely in terms of Lie algebras, but as pointed out in [97] the problem seems to be substantially harder than for $\ell \neq 2$.

It is now easy to deduce theorem 1.5.1 (1):

Proof. The proof follows closely that of theorem 1.4.2 (1): we can replace G first by $H = G\cdot(1+8\mathbb{Z}_2)$ and then by $H_0 = H \cap \mathrm{SL}_2(\mathbb{Z}_2)$ without altering $L(G)$ nor G', so we are reduced to working with subgroups of $\mathrm{SL}_2(\mathbb{Z}_2)$. Note now that $n \geq 2$ since by hypothesis every element in G (and hence in H_0) has its off-diagonal coefficients divisible by 4. Theorem 1.5.2 then guarantees that H_0 contains $\mathcal{B}_2(6n)$, so $G' = H_0'$ contains $\mathcal{B}_2(12n+2)$ because of lemma 1.3.3. □

1.6 Lie algebras modulo ℓ^n

Fix any prime number ℓ and let L be a topologically open and closed, $\mathbb{Z}_\ell$-Lie subalgebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. The same arguments of the previous section, namely an application of lemma 1.3.11, yield the existence of a basis of L of the form

$$x_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix}, x_2 = \begin{pmatrix} b_{11} & b_{12} \\ 0 & -b_{11} \end{pmatrix}, x_3 = \begin{pmatrix} 0 & c_{12} \\ 0 & 0 \end{pmatrix}.$$

Definition 1.6.1. A basis of this form will be called a *reduced* basis.

There is clearly no uniqueness of such an object, but in what follows we will just assume that the choice of a reduced basis has been made.

Notation. We let $k(L)$, or simply k, denote the number $\min_{m\in L} v_\ell(m_{21})$, where m_{21} is the bottom-left coefficient of m in the standard matrix representation of elements of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Furthermore, for every positive n we denote by $L\left(\ell^n\right)$ be the image of the mod-ℓ^n reduction map $\pi_n : L \to \mathfrak{sl}_2(\mathbb{Z}/\ell^n\mathbb{Z})$; clearly $L\left(\ell^n\right)$ is a Lie algebra over $\mathbb{Z}/\ell^n\mathbb{Z}$.

Remark 1.6.2. It is apparent from the very definition of a reduced basis that $k(L) = v_\ell(a_{21})$. Also notice that, by definition, the images of x_1, x_2, x_3 in $L\left(\ell^n\right)$ generate it as a $(\mathbb{Z}/\ell^n\mathbb{Z})$-module.

The following statement allows us to deduce properties of $G(\ell^n)$ from corresponding properties of $L(\ell^n)$:

Proposition 1.6.3. *Suppose L as above is obtained as $\overline{\Theta(G)}$ for a certain closed subgroup G of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (whose reduction modulo 2 is trivial if $\ell = 2$). For every integer $m \geq 1$ let $G(\ell^m)$ be the image of G in $\mathrm{GL}_2(\mathbb{Z}/\ell^m\mathbb{Z})$, and let $j_m = \left|\{i \in \{1,2,3\} \mid x_i \not\equiv 0 \pmod{\ell^m}\}\right|$ (that is, exactly j_m among x_1, x_2 and x_3 are nonzero modulo ℓ^m). For every $n \geq 1$ the following are the only possibilities (recall that $v = v_\ell(2)$):*

- *j_n is at most 1 and $G(\ell^n)$ is abelian.*
- *$j_n = 2$ and either $j_{2n} = 3$ or $G(\ell^{n-k(L)+1-2v})$ is contained in the subgroup of upper-triangular matrices (up to a change of coordinates in $\mathrm{GL}_2(\mathbb{Z}_\ell)$).*
- *$j_n = 3$ and L contains $\ell^{n+2k(L)-1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.*

Remark 1.6.4. The exponent $n + 2k(L) - 1$ is best possible: fix integers $k \geq 0$, $n \geq 1$ and let L be the Lie algebra generated (as a $\mathbb{Z}_\ell$-module) by $x_1 = \begin{pmatrix} 1 & 0 \\ \ell^k & -1 \end{pmatrix}, x_2 = \begin{pmatrix} \ell^{k+n-1} & 0 \\ 0 & -\ell^{k+n-1} \end{pmatrix}$, and $x_3 = \begin{pmatrix} 0 & \ell^{n-1} \\ 0 & 0 \end{pmatrix}$. Then clearly $k(L) = k$, $j_n(L) = 3$, and it is easy to check that $n + 2k - 1$ is the smallest exponent s such that $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$ is contained in L.

Proof. Assume first $j_n \leq 1$. It is clear that every element of $G(\ell^n)$ can we written as $\lambda \operatorname{Id} + m_n$ for some $\lambda \in \mathbb{Z}/\ell^n\mathbb{Z}$ and $m_n \in L(\ell^n)$. Now L is generated by x_1, x_2, x_3, so in turn every m_n is of the form $\pi_n(\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3)$, and since at most one of $\pi_n(x_1), \pi_n(x_2), \pi_n(x_3)$ is non-zero we can find an $l_n \in L(\ell^n)$ such that, for every m_n, there exists a scalar $\mu \in \mathbb{Z}/\ell^n\mathbb{Z}$ with $m_n = \mu\, l_n$. It follows that every element of $G(\ell^n)$ can be written as $\lambda \operatorname{Id} + \mu\, l_n$ for suitable λ, μ, and since Id and l_n commute our claim follows.

Next consider the case $j_n = 2$. We can safely assume that $j_{2n} = 2$, for otherwise we are done (notice that $j_{2n} \geq j_n = 2$). Under this assumption, it is clear that for $i = 1, 2, 3$ we have $\pi_n(x_i) = 0$ if and only if $\pi_{2n}(x_i) = 0$. Suppose first $\pi_n(x_1) = 0$, so that $k(L) \geq 1$. Then $G(\ell^n)$ is a subset of

$$\mathbb{Z}/\ell^n\mathbb{Z} \cdot \operatorname{Id} + \mathbb{Z}/\ell^n\mathbb{Z} \cdot \pi_n(x_2) + \mathbb{Z}/\ell^n\mathbb{Z} \cdot \pi_n(x_3),$$

and $\operatorname{Id}, \pi_n(x_2), \pi_n(x_3)$ are upper-triangular matrices, so $G(\ell^n)$ – hence also $G(\ell^{n-k(L)+1-2v})$, since $k(L) \geq 1$ – is in triangular form.
Suppose next $\pi_n(x_1) \neq 0$. Assume that $\pi_n(x_3) = 0$ (the other case being analogous, as we are only going to use that x_2 is upper triangular). L is a Lie algebra, hence so is $L\left(\ell^{2n}\right)$; furthermore, every element in $L\left(\ell^{2n}\right)$ is a combination of $\pi_{2n}(x_1), \pi_{2n}(x_2)$ with coefficients in $\mathbb{Z}/\ell^{2n}\mathbb{Z}$. In particular, there exist $\xi_1, \xi_2 \in \mathbb{Z}/\ell^{2n}\mathbb{Z}$ such that

$$\begin{aligned}[x_1, x_2] - 2b_{11}x_1 + 2a_{11}x_2 &= \begin{pmatrix} -a_{21}b_{12} & 4(a_{11}b_{12} - a_{12}b_{11}) \\ 0 & a_{21}b_{12} \end{pmatrix} \\ &\equiv \xi_1 x_1 + \xi_2 x_2 \pmod{\ell^{2n}}.\end{aligned}$$

Matching the bottom-left coefficients we find $\xi_1 a_{21} \equiv 0 \pmod{\ell^{2n}}$, so, using $v_\ell(a_{21}) = k(L)$, we immediately deduce $\xi_1 \equiv 0 \pmod{\ell^{2n-k(L)}}$. Reducing the above congruence modulo $\ell^{2n-k(L)}$ we then have the relations

$$\begin{cases} -a_{21}b_{12} \equiv \xi_2 b_{11} \pmod{\ell^{2n-k(L)}} \\ 4(a_{11}b_{12} - a_{12}b_{11}) \equiv \xi_2 b_{12} \pmod{\ell^{2n-k(L)}}. \end{cases} \tag{1.2}$$

We now introduce the vector $y = \begin{pmatrix} b_{12} \\ -2b_{11} \end{pmatrix} \in \mathbb{Z}_\ell^2$. An immediate calculation shows that this is an exact eigenvector for x_2 (associated with the eigenvalue $-b_{11}$), and on the other hand it is also an approximate eigenvector for $2x_1$, in the sense that $2x_1 \cdot y \equiv (\xi_2 - 2a_{11})\, y \pmod{\ell^{2n-k(L)}}$. Indeed,

$$2x_1 \cdot y = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix} \begin{pmatrix} 2b_{12} \\ -4b_{11} \end{pmatrix} = \begin{pmatrix} 2a_{11}b_{12} - 4a_{12}b_{11} \\ 2a_{21}b_{12} + 4a_{11}b_{11} \end{pmatrix},$$

and using (1.2) we find

$$\begin{aligned} 2x_1 \cdot y &= \begin{pmatrix} 2a_{11}b_{12} - 4a_{12}b_{11} \\ 2a_{21}b_{12} + 4a_{11}b_{11} \end{pmatrix} \\ &\equiv \begin{pmatrix} 2a_{11}b_{12} + \xi_2 b_{12} - 4a_{11}b_{12} \\ -2\xi_2 b_{11} + 4a_{11}b_{11} \end{pmatrix} \\ &\equiv (\xi_2 - 2a_{11})y \pmod{\ell^{2n-k(L)}} \end{aligned}$$

as claimed.

Now if $\ell \neq 2$ we immediately deduce $x_1 \cdot y \equiv \left(\frac{\xi_2}{2} - a_{11}\right) y \pmod{\ell^{2n-k(L)}}$. If, on the other hand, $\ell = 2$, then we would like to prove that $v_2(\xi_2) \geq 1$ in order to be able to divide by 2. Observe that y is not zero modulo 2^{n+1}, since its coordinates are (up to a factor of 2) the entries of x_2, which we have assumed not to reduce to zero in $L(2^n)$.
Let $\alpha = \min\{v_2(2b_{11}), v_2(b_{21})\} \leq n$ and reduce the last congruence modulo $2^{\alpha+1}$. Then we have $2x_1 \cdot y \equiv x_1 \cdot (2y) \equiv 0 \pmod{2^{\alpha+1}}$, so $(\xi_2 - 2a_{11})\, y \equiv 0 \pmod{2^{\alpha+1}}$, which implies that ξ_2 is even (that is to say, $v_2(\xi_2) \geq 1$), for otherwise multiplying by $\lambda - 2a_{11}$ would be invertible modulo $2^{\alpha+1}$ and we would find $y \equiv 0 \pmod{2^{\alpha+1}}$, contradicting the definition of α. It follows that we can indeed divide the above congruence by 2 to get

$$x_1 \cdot y \equiv \left(\frac{\xi_2}{2} - a_{11}\right) y \pmod{2^{2n-k(L)-1}}.$$

Equivalently, the following congruence holds for *every* prime ℓ:

$$x_1 \cdot y \equiv \left(\frac{\xi_2}{2} - a_{11}\right) y \pmod{\ell^{2n-k(L)-v}}.$$

Note now that it is in fact true for *every* ℓ that y is not zero modulo ℓ^{n+v} (its coordinates are, up to a factor of 2, the entries of x_2, which we have assumed not to reduce to zero modulo ℓ^n).
Let again $\alpha = \min\{v_\ell(2b_{11}), v_\ell(b_{21})\} \leq n - 1 + v$ and set $\tilde{y} = \ell^{-\alpha} y$. Dividing by ℓ^α the congruence $x_1 \cdot y \equiv \left(\frac{\xi_2}{2} - a_{11}\right) y \pmod{\ell^{2n-k(L)-v}}$ we get $x_1 \cdot \tilde{y} \equiv \left(\frac{\xi_2}{2} - a_{11}\right) \tilde{y} \pmod{\ell^{n-k(L)+1-2v}}$, where $\tilde{y} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}$ is a vector at least one of whose coordinates is an ℓ-adic unit. Assume by symmetry that $v_\ell(\tilde{y}_1) = 0$ and introduce the base-change matrix $P = \begin{pmatrix} \tilde{y}_1 & 0 \\ \tilde{y}_2 & 1 \end{pmatrix}$: this is then an element of $\mathrm{GL}_2(\mathbb{Z}_\ell)$, since its determinant $\tilde{y}_1$ is not divisible by ℓ.
An element of $G(\ell^{n-k(L)+1-2v})$ will be of the form $g = \lambda \,\mathrm{Id} + \mu_1 x_1 + \mu_2 x_2$, so by construction conjugating G via P puts $G(\ell^{n-k(L)+1-2v})$ in upper-triangular form. Indeed, the first column of x_i (for $i = 1, 2$) in the coordinates defined by P is given by

$$\begin{aligned} P^{-1} x_i P \begin{pmatrix} 1 \\ 0 \end{pmatrix} = P^{-1} x_i \cdot \tilde{y} &= P^{-1} \left((\xi_2/2 - a_{11}) \tilde{y} + \ell^{n-k(L)+1-2v} w \right) \\ &= (\xi_2/2 - a_{11}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \ell^{n-k(L)+1-2v} P^{-1} w \\ &\equiv (\xi_2/2 - a_{11}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \pmod{\ell^{n-k(L)+1-2v}} \end{aligned}$$

where w is a suitable vector in $\mathbb{Z}_\ell^2$ (that vanishes for $i = 2$).

Finally, suppose $j_n = 3$. Then we have in particular $\pi_n(x_3) \neq 0$, so $v_\ell(c_{12}) \leq n - 1$. As L is a Lie algebra, we see that it contains

$$x_4 = [x_1, x_3] - 2a_{11} x_3 = \begin{pmatrix} -a_{21} c_{12} & 0 \\ 0 & a_{21} c_{12} \end{pmatrix},$$

whose diagonal entries have valuation at most $v_\ell(a_{21}) + v_\ell(c_{12}) \leq k(L) + (n-1)$. Furthermore, L also contains the linear combination

$$x_5 = \ell^{n+k(L)-1}x_1 + \frac{\ell^{n+k(L)-1}a_{11}}{a_{21}c_{12}}x_4 - \frac{\ell^{n+k(L)-1}a_{12}}{c_{12}}x_3 = \begin{pmatrix} 0 & 0 \\ \ell^{n+k(L)-1}a_{21} & 0 \end{pmatrix} :$$

notice that the coefficients $\dfrac{\ell^{n+k(L)-1}a_{11}}{a_{21}c_{12}}$ and $\dfrac{\ell^{n+k(L)-1}a_{12}}{c_{12}}$ have positive ℓ-adic valuation by what we have already shown, and that the valuation of the only non-zero coefficient of x_5 is $n+2k(L)-1$. Setting

$$s_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, s_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, s_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

we see that L contains the three elements $x_3 = c_{12}s_1$, $x_4 = -a_{21}c_{12}s_2$, $x_5 = \ell^{n+k(L)-1}a_{21}s_3$. By what we have already proved we have

$$\max\left\{v_\ell(c_{12}), v_\ell(-a_{21}c_{12}), v_\ell\left(\ell^{n+k(L)-1}a_{21}\right)\right\} = n + 2k(L) - 1,$$

so the $\mathbb{Z}_\ell$-module generated by x_3, x_4, x_5 contains $\ell^{n+2k(L)-1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, and a fortiori so does L. □

Corollary 1.6.5. *Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ satisfying property $(\star\star)$ of theorem 1.4.2 (resp. $G(4) = \{\mathrm{Id}\}$ and $\det(G) \equiv 1 \pmod 8$ if $\ell = 2$). Then for every positive integer $n \geq k(L(G))$ at least one of the following holds:*

1. *$G(\ell^n)$ is abelian.*
2. *$G(\ell^{n-k(L(G))+1-2v})$ is contained in the subgroup of upper-triangular matrices (up to a change of coordinates in $\mathrm{GL}_2(\mathbb{Z}_\ell)$).*
3. *G' contains the principal congruence subgroup*

$$\mathcal{B}_\ell(16n-4) = \left(\mathrm{Id} + \ell^{16n-4}\mathfrak{gl}_2(\mathbb{Z}_\ell)\right) \cap \mathrm{SL}_2(\mathbb{Z}_\ell),$$

if ℓ is odd, and it contains $\mathcal{B}_2(48n-10)$, if $\ell = 2$.

Proof. To ease the notation set $L = L(G)$. Consider $L(\ell^n)$ and distinguish cases depending on j_n as in the statement of the previous proposition. If $j_n \leq 1$ we are in case (1) and we are done. If $j_n \geq 2$ we begin by proving that either (2) holds or L contains $\ell^{4n-1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.
If $j_n = 2$ and $j_{2n} = 2$, then we are in situation (2) by the previous proposition. If, on the other hand, $j_n = 2$ and $j_{2n} = 3$, then (again by proposition 1.6.3) we have

$$L \supseteq \ell^{2n+2k(L)-1}\mathfrak{sl}_2(\mathbb{Z}_\ell) \supseteq \ell^{4n-1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$$

since $n \geq k(L)$. Finally, for $j_n = 3$ the proposition yields directly

$$L \supseteq \ell^{n+2k(L)-1}\mathfrak{sl}_2(\mathbb{Z}_\ell) \supseteq \ell^{3n-1}\mathfrak{sl}_2(\mathbb{Z}_\ell).$$

In all cases, property $(\star\star)$ (resp. theorem 1.5.1 (1) for $\ell = 2$) now implies that G' contains $\mathcal{B}_\ell(16n-4)$ (resp. $\mathcal{B}_2(48n-10)$) as claimed. □

1.7 Application to Galois groups

We now plan to apply the above machinery to the Galois representations attached to an elliptic curve. Let therefore K be a number field and E an elliptic curve over K without (potential) complex multiplication.

Notation. ℓ is any rational prime, n a positive integer and G_ℓ the image of $\mathrm{Gal}(\overline{K}/K)$ inside $\mathrm{Aut}\, T_\ell(E) \cong \mathrm{GL}_2(\mathbb{Z}_\ell)$. As before, v is 0 or 1 according to whether ℓ is respectively odd or even.

If ℓ is odd (resp. $\ell = 2$), then by theorem 1.4.2 (resp. theorem 1.5.1) we know that either G_ℓ contains a subgroup H_ℓ satisfying $[G_\ell : H_\ell] \leq 24$ (respectively $[G_\ell : H_\ell] \leq 192$ for $\ell = 2$) and the hypotheses of corollary 1.6.5, or otherwise $G'_\ell = \mathrm{SL}_2(\mathbb{Z}_\ell)$. In this second case we put $H_\ell = G_\ell$.
We also denote K_ℓ the extension of K fixed by H_ℓ. The degree $[K_\ell : K]$ is then bounded by 24, for odd ℓ, and $2 \cdot |\mathrm{GL}_2(\mathbb{Z}/4\mathbb{Z})| = 2 \cdot 96$, for $\ell = 2$. For a fixed ℓ, upon replacing K with K_ℓ we are reduced to the case where G_ℓ satisfies the hypotheses of corollary 1.6.5. In order to apply this result we want to have numerical criteria to exclude the 'bad' cases (1) and (2). These numerical bounds form the subject of lemma 1.7.1 and proposition 1.7.4 below, whose proofs are inspired by the arguments of [71] and [69].

Lemma 1.7.1. *Suppose E/K does not admit potential complex multiplication. If $\ell^n \nmid b_0(K, E)$ the group $G_\ell(\ell^n)$ cannot be put in triangular form.*

Proof. Suppose that $G_\ell(\ell^n)$ is contained (up to a change of basis) in the group of upper-triangular matrices. The subgroup Γ of $E[\ell^n]$ given (in the coordinates in which $G_\ell(\ell^n)$ is triangular) by

$$\Gamma = \left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} \mid a \in \mathbb{Z}/\ell^n\mathbb{Z} \right\}$$

is $\mathrm{Gal}(\overline{K}/K)$-stable, hence defined over K. Consider then $E^* = E/\Gamma$ and the natural projection $\pi : E \to E^*$ of degree $|\Gamma| = \ell^n$. By theorem 1.2.8 we also have an isogeny $E^* \to E$ of degree b, with $b \mid b_0(K, E)$. Composing the two we get an endomorphism of E that kills Γ, and therefore corresponds (since $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is annihilated by ℓ^n) to multiplication by a certain $\ell^n d$, $d \in \mathbb{Z}$. Taking degrees we get $\ell^n \cdot b = |\Gamma| \cdot b = d^2 \ell^{2n}$, so $\ell^n \mid b$ and $\ell^n \mid b_0(K, E)$. □

Corollary 1.7.2. *Let L be the special Lie algebra of G_ℓ (supposing that $G_\ell(2)$ is trivial if $\ell = 2$). The inequality $k(L) \leq v_\ell(b_0(K, E))$ holds, so that in particular $\ell^{k(L)} \mid b_0(K, E)$.*

Proof. Let $t = v_\ell(b_0(K, E))$. If by contradiction we had $k(L) \geq t + 1$, then $L\left(\ell^{t+1}\right)$ would be triangular, and therefore so would be $G_\ell(\ell^{t+1}) \subseteq \mathbb{Z}/\ell^{t+1}\mathbb{Z} \cdot \mathrm{Id} + L\left(\ell^{t+1}\right)$, which is absurd, since $\ell^{t+1} \nmid b_0(K, E)$. □

Corollary 1.7.3. *If $\ell^n \nmid b_0(K, E)$ the group $G_\ell(\ell^n)$ does not consist entirely of scalar matrices. In particular this is true for $G_\ell(\ell^{v_\ell(b_0(K,E))+1})$.*

Using this last corollary we find:

Proposition 1.7.4. *If ℓ^{2n} does not divide $b_0(K, E)^4 b_0(K, E \times E)$ the group $G_\ell(\ell^n)$ is not abelian. In particular, the group $G_\ell(\ell)$ is not abelian if ℓ does not divide $b_0(K, E) b_0(K, E \times E)$.*

Proof. For the sake of simplicity set $d = b_0(K, E)$. By the previous corollary, there is an $\alpha \in G_\ell$ whose image modulo $\ell^{1+v_\ell(d)}$ is not a scalar matrix. Suppose now that $G_\ell(\ell^n)$ is abelian. Consider the subgroup $\Gamma = \{(x, \alpha(x)) \mid x \in E[\ell^n]\} \subset E \times E$; this is defined over K, since for any $\gamma \in G_\ell(\ell^n)$ we have $\gamma \cdot (x, \alpha(x)) = (\gamma \cdot x, \gamma \cdot \alpha(x)) = (\gamma \cdot x, \alpha(\gamma \cdot x))$ as $G_\ell(\ell^n)$ is commutative. We can therefore form the quotient K-variety $E^* = (E \times E)/\Gamma$, which comes equipped with a natural isogeny $E \times E \twoheadrightarrow E^*$ of degree $|\Gamma| = E[\ell^n] = \ell^{2n}$; on the other hand, theorem 1.2.8 yields the existence of a K-isogeny $E^* \to E \times E$ of degree $b \mid b_0(K, E \times E)$. Composing the two we obtain an endomorphism ψ of $E \times E$, which (given that E does not admit complex multiplication) can be represented as a 2×2 matrix $\begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}$ with coefficients in $\mathbb{Z}$ and nonzero determinant.
Now since ψ kills Γ we must have $e_{11}x + e_{12}\alpha(x) = 0$ and $e_{21}x + e_{22}\alpha(x) = 0$ for every $x \in E[\ell^n]$. Let $\eta = \min\{v_\ell(e_{ij})\}$ and suppose by contradiction $\eta < n - v_\ell(d)$. For the sake of simplicity, let us assume this minimum is attained for e_{12} (the other cases being completely analogous: the situation is manifestly symmetric in the index i, and to show that it is symmetric in j it is enough to compose with α^{-1}, which is again a non-scalar matrix). Dividing the equation $e_{11}x + e_{12}\alpha(x) = 0$ by ℓ^η we get

$$\frac{e_{11}}{\ell^\eta}x + \frac{e_{12}}{\ell^\eta}\alpha(x) \equiv 0 \pmod{\ell^{n-\eta}} \quad \forall x \in E[\ell^n],$$

whence

$$\frac{e_{11}}{\ell^\eta}x + \frac{e_{12}}{\ell^\eta}\alpha(x) = 0 \quad \forall x \in E[\ell^{n-\eta}],$$

where now $\frac{e_{12}}{\ell^\eta}$ is invertible modulo $\ell^{n-\eta}$, being relatively prime to ℓ. Multiplying by the inverse of $\frac{e_{12}}{\ell^\eta}$, then, we find that

$$\alpha(x) = -\frac{e_{11}}{\ell^\eta}\left(\frac{e_{12}}{\ell^\eta}\right)^{-1} x \quad \forall x \in E[\ell^{n-\eta}],$$

i.e. α is a scalar modulo $\ell^{n-\eta}$. By definition of α, this implies $\ell^{n-\eta} \mid d$, so $n - \eta \leq v_\ell(d)$, a contradiction. It follows that $\ell^{2n}\ell^{-2v_\ell(d)} \mid \ell^{2\eta} \mid \det\begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}$. Squaring this last divisibility we find

$$\ell^{4n}\ell^{-4v_\ell(d)} \mid \left(\det\begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}\right)^2 = \deg(\psi) = b\ell^{2n},$$

so $\ell^{2n}\ell^{-4v_\ell(d)} \mid b$ and $\ell^{2n} \mid \ell^{4v_\ell(d)} b_0(K, E \times E) \mid d^4\, b_0(K, E \times E)$. The second assertion follows immediately from the fact that ℓ is prime. □

With these results at hand it is now immediate to deduce the following theorem, where we use the notation introduced at the beginning of this section and the symbol $\mathcal{B}_\ell(n)$ of section 1.3.

Theorem 1.7.5. *Let ℓ be a prime and set $D(\ell) = b_0(K_\ell, E)^5 b_0(K_\ell, E \times E)$. Let n be a positive integer. Suppose that ℓ^{n-v} does not divide $D(\ell)$: then H'_ℓ contains $\mathcal{B}_\ell(16n - 4)$, for odd ℓ, and it contains $\mathcal{B}_2(48n - 10)$, for $\ell = 2$.*

Proof. By the discussion at the beginning of this section there are two possibilities: if the derived subgroup G'_ℓ is all of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ then the conclusion is obvious since $H_\ell = G_\ell$; if this is not the case,

then H_ℓ satisfies the hypotheses of corollary 1.6.5. Note that the image of $\mathrm{Gal}(\overline{K_\ell}/K_\ell)$ in $\mathrm{Aut}\, T_\ell(E)$ is exactly H_ℓ by construction. We wish to apply corollary 1.6.5 to $G = H_\ell$, assuming that ℓ^{n-v} does not divide $D(\ell)$.
Since $\ell^{k(L)} \mid b_0(K_\ell, E)$ by corollary 1.7.2, we deduce $\ell^{n-k(L)-v} \nmid b_0(K_\ell, E)^4 b_0(K_\ell, E \times E)$, and a fortiori $\ell^{n-k(L)+1-2v} \nmid b_0(K_\ell, E)^4 b_0(K_\ell, E \times E)$. Lemma 1.7.1 then implies that $G(\ell^{n-k(L)+1-2v})$ cannot be put in triangular form, and on the other hand $\ell^{n-v} \nmid b_0(K_\ell, E)^5 b_0(K_\ell, E \times E)$ implies that ℓ^{2n} does not divide $b_0(K_\ell, E)^4 b_0(K_\ell, E \times E)$, so $G(\ell^n)$ is not abelian (thanks to proposition 1.7.4). It then follows from corollary 1.6.5 that $G' = H'_\ell$ contains the principal congruence subgroup $\mathcal{B}_\ell(16n-4)$ (resp. $\mathcal{B}_\ell(48n-10)$ for $\ell = 2$). □

Corollary 1.7.6. *Let $D(\infty) = b_0(K, E; 120)^5 b_0(K, E \times E; 120)$ and ℓ be an odd prime. If ℓ^n does not divide $D(\infty)$, then H'_ℓ contains $\mathcal{B}_\ell(16n-4)$.*

Proof. As $[K_\ell : K] \leq 120$ we find that

$$D(\ell) = b_0(K_\ell, E)^5 b_0(K_\ell, E \times E) \mid b_0(K, E; 120)^5 b_0(K, E \times E; 120) = D(\infty),$$

so the result follows from the theorem since $\ell^n \nmid D(\infty)$ implies $\ell^n \nmid D(\ell)$. □

Corollary 1.7.7. *Notation as above. The index $[\mathrm{SL}_2(\mathbb{Z}_\ell) : (H'_\ell \cap \mathcal{B}_\ell(1))]$ can be written as the product $|\mathrm{SL}_2(\mathbb{F}_\ell)| B(\ell)$, where for $\ell \neq 2$ the number $B(\ell)$ is a power of ℓ dividing $\ell^{33} \cdot D(\ell)^{48}$ (respectively $B(2)$ is a power of 2 dividing $2^{255} D(2)^{144}$).*

Proof. We can write the index $[\mathrm{SL}_2(\mathbb{Z}_\ell) : (H'_\ell \cap \mathcal{B}_\ell(1))]$ as

$$[\mathrm{SL}_2(\mathbb{Z}_\ell) : \mathcal{B}_\ell(1)] \cdot [\mathcal{B}_\ell(1) : (H'_\ell \cap \mathcal{B}_\ell(1))] = |\mathrm{SL}_2(\mathbb{F}_\ell)| \cdot [\mathcal{B}_\ell(1) : (H'_\ell \cap \mathcal{B}_\ell(1))],$$

so we just need to prove that $B(\ell) = [\mathcal{B}_\ell(1) : (H'_\ell \cap \mathcal{B}_\ell(1))]$ divides $\ell^{33} D(\ell)^{48}$ (and the analogous statement for $\ell = 2$). Notice that since $\mathcal{B}_\ell(1)$ is a pro-ℓ group the number $B(\ell)$ is a power of ℓ.
Choose n such that $\ell^{n-v} \mid\mid D(\ell)$: then $\ell^{n+1-v} \nmid D(\ell)$, and therefore the above theorem implies that H'_ℓ contains $\mathcal{B}_\ell(16(n+1)-4) \subseteq \mathcal{B}_\ell(1)$ (resp. $\mathcal{B}_2(48(n+1)-10)$ for $\ell = 2$): the index of $\mathcal{B}_\ell(16(n+1)-4)$ in $\mathcal{B}_\ell(1)$ is $\ell^{3(16(n+1)-5)}$, so we get

$$[\mathcal{B}_\ell(1) : (H'_\ell \cap \mathcal{B}_\ell(1))] \mid \ell^{48n+33} \mid \ell^{33} \cdot D(\ell)^{48}$$

for $\ell \neq 2$, and likewise we have

$$[\mathcal{B}_2(1) : (H'_2 \cap \mathcal{B}_2(1))] \mid 2^{3(48(n-1)+85)} \mid 2^{255} D(2)^{144}$$

for $\ell = 2$. □

1.8 The determinant and the large primes

We now turn to studying the determinant of the adelic representation and the behaviour at the very large primes.

Proposition 1.8.1. *The index*

$$\left[\widehat{\mathbb{Z}}^\times : \prod_\ell \det \rho_\ell(\mathrm{Gal}(\overline{K}/K))\right]$$

is bounded by $[K : \mathbb{Q}]$.

Proof. The Weil pairing induces an identification of the determinant $\mathrm{Gal}(\overline{K}/K) \xrightarrow{\rho_\ell} G_\ell \xrightarrow{\det} \mathbb{Z}_\ell^\times$ with $\mathrm{Gal}(\overline{K}/K) \xrightarrow{\chi_\ell} \mathbb{Z}_\ell^\times$, where χ_ℓ denotes the ℓ-adic cyclotomic character; by Galois theory we have

$$\prod_\ell \det \rho_\ell \left(\mathrm{Gal}(\overline{K}/K)\right) = \prod_\ell \chi_\ell \left(\mathrm{Gal}(\overline{K}/K)\right) \cong \mathrm{Gal}\left(K\left(\mu_\infty\right)/K\right).$$

Let $F = K \cap \mathbb{Q}\left(\mu_\infty\right)$: it is a finite Galois extension of $\mathbb{Q}$. As $\mathbb{Q}\left(\mu_\infty\right)$ is Galois over $\mathbb{Q}$, the restriction map $\mathrm{Gal}\left(K\left(\mu_\infty\right)/K\right) \to \mathrm{Gal}\left(\mathbb{Q}\left(\mu_\infty\right)/F\right)$ is well-defined and induces an isomorphism. Therefore

$$\begin{aligned}\left[\widehat{\mathbb{Z}}^\times : \prod_\ell \chi_\ell(\mathrm{Gal}(\overline{K}/K))\right] &= \left[\mathrm{Gal}\left(\mathbb{Q}\left(\mu_\infty\right)/\mathbb{Q}\right) : \mathrm{Gal}\left(\mathbb{Q}\left(\mu_\infty\right)/F\right)\right] \\ &= [F:\mathbb{Q}] \leq [K:\mathbb{Q}]\end{aligned}$$

as claimed. □

We will also need a surjectivity result (on SL_2) modulo ℓ for every ℓ sufficiently large: as previously mentioned, these are essentially the ideas of [71] and [68], in turn inspired by those of Serre.

Lemma 1.8.2. *If $\ell \nmid b_0(K, E \times E; 2) b_0(K, E; 60)$ then the group $G_\ell(\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$.*

Proof. Let ℓ be a prime for which $G_\ell(\ell)$ does not contain $\mathrm{SL}_2(\mathbb{F}_\ell)$ and let, for the sake of clarity, $G = G_\ell(\ell)$. By theorem 1.3.13, if G does not contain $\mathrm{SL}_2(\mathbb{F}_\ell)$, then the following are the only possibilities:

1. G is contained in a Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$: by definition, such a subgroup fixes a line, therefore $\ell \mid b_0(K, E)$ by lemma 1.7.1.

2. G is contained in the normalizer of a Cartan subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$: let $\mathcal{C}$ be this Cartan subgroup and N its normalizer. By Dickson's classification $\mathcal{C}$ has index 2 in N, so the morphism $\mathrm{Gal}(\overline{K}/K) \to G \to \dfrac{G}{G \cap \mathcal{C}} \hookrightarrow \dfrac{N}{\mathcal{C}}$ induces a quadratic character of $\mathrm{Gal}(\overline{K}/K)$, whose kernel corresponds to a certain field K' satisfying $[K' : K] \leq |N/\mathcal{C}| = 2$. By construction, the image of $\mathrm{Gal}(\overline{K'}/K')$ in $\mathrm{Aut}\left(E[\ell]\right)$ is contained in $\mathcal{C}$, so applying proposition 1.7.4 to $E_{K'}$ we get

   $$\ell \mid b_0(K', E) b_0(K', E \times E) \mid b_0(K, E; 2) b_0(K, E \times E; 2).$$

   Notice that this also covers the case of G being contained in a Cartan subgroup.

3. The projectivization $\mathbb{P}G$ of G is a finite group of order at most 60: we essentially copy the previous argument. Let $H = \mathbb{P}G$; then we have a morphism

   $$\mathrm{Gal}(\overline{K}/K) \to G \to \frac{\mathbb{F}_\ell^\times G}{\mathbb{F}_\ell^\times} = H$$

   whose kernel defines an extension K'' of K with $[K'' : K] = |H| \leq 60$ and such that the image of the representation of $\mathrm{Gal}\left(\overline{K''}/K''\right)$ on $E[\ell]$ is contained in $\mathbb{F}_\ell^\times$: lemma 1.7.1 then yields $\ell \mid b_0(K'', E) \mid b_0(K, E; 60)$.

It is then apparent that the lemma is true with the condition

$$\ell \nmid b_0(K, E) b_0(K, E \times E) b_0(K, E; 2) b(K, E \times E; 2) b_0(K, E; 60);$$

however, since

$$b_0(K, E) \mid b_0(K, E; 2) \mid b_0(K, E; 60), \quad b_0(K, E \times E) \mid b_0(K, E \times E; 2),$$

and since ℓ is prime, we see that ℓ divides

$$b_0(K,E)b_0(K,E\times E)b_0(K,E;2)b_0(K,E\times E;2)b_0(K,E;60)$$

if and only if it divides $b(K,E\times E;2)b_0(K,E;60)$, which finishes the proof. □

Corollary 1.8.3. *Let $\Psi = 30\cdot b_0(K,E\times E;2)b_0(K,E;60)$. If $\ell\nmid\Psi$, then G'_ℓ is all of $\mathrm{SL}_2(\mathbb{Z}_\ell)$.*

Proof. The previous lemma implies that $G_\ell(\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$, and by hypothesis ℓ is strictly larger than 3, so the corollary follows from lemma 1.3.15. □

1.9 The adelic index and some consequences

We have thus acquired a good understanding of the ℓ-adic representation for every prime ℓ, and we are now left with the task of bounding the overall index of the full adelic representation. The statement we are aiming for is:

Theorem 1.9.1. *Let E/K be an elliptic curve without complex multiplication with stable Faltings height $h(E)$. Let $\rho_\infty : \mathrm{Gal}(\overline{K}/K)\to \mathrm{GL}_2\left(\widehat{\mathbb{Z}}\right)$ be the adelic Galois representation associated with E, and set*

$$\Psi = 2\cdot 3\cdot 5\cdot b_0(K,E\times E;2)b_0(K,E;60),\quad D(\infty) = b_0(K,E;24)^5 b_0(K,E\times E;24);$$

let moreover K_2 be as in section 1.7 and

$$D(2) = b_0(K_2,E)^5 b_0(K_2,E\times E).$$

With this notation we have

$$\left[\mathrm{GL}_2\left(\widehat{\mathbb{Z}}\right) : \rho_\infty \mathrm{Gal}(\overline{K}/K)\right] \le [K:\mathbb{Q}]\cdot 2^{222}\cdot D(2)^{144}\cdot \mathrm{rad}(\Psi)^{36}\cdot D(\infty)^{48},$$

where $\mathrm{rad}(\Psi) = \prod_{\ell\mid\Psi}\ell$ is the product of the primes dividing Ψ.

The strategy of proof, which essentially goes back to Serre, is to pass to a suitable extension of K over which the adelic representation decomposes as a direct product and then use the previous bounds. For this we will need some preliminaries. If L is any number field, we let $L_{cyc} = L\left(\mu_\infty\right)$ be its maximal cyclotomic extension. From the exact sequence

$$1\to \frac{\mathrm{SL}_2(\widehat{\mathbb{Z}})}{\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)}\to \frac{\mathrm{GL}_2(\widehat{\mathbb{Z}})}{\rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)}\to \frac{\widehat{\mathbb{Z}}^\times}{\det\circ\rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)}\to 1$$

we see that $[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)]$ equals

$$[\widehat{\mathbb{Z}}^\times : \det\circ\rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)]\cdot[\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)],$$

where the first term is bounded by $[K:\mathbb{Q}]$ thanks to proposition 1.8.1. It thus remains to understand the term $[\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)]$. Let $\mathcal{P}$ be the (finite) set consisting of $2,3,5$, and the prime numbers ℓ for which G_ℓ does not contain $\mathrm{SL}_2(\mathbb{Z}_\ell)$, and let F be the field generated over K by $\bigcup_{\ell\in\mathcal{P}} E[\ell]$. It is clear that

$$[\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)] \le [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/F_{cyc}\right)\right)].$$

Notation. We set $S = \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/F_{cyc}\right)\right) \subseteq \mathrm{SL}_2(\widehat{\mathbb{Z}}) = \prod_\ell \mathrm{SL}_2(\mathbb{Z}_\ell)$ and let S_ℓ be the projection of S on $\mathrm{SL}_2(\mathbb{Z}_\ell)$.

The core of the argument is contained in the following proposition.

Proposition 1.9.2. *Let $B(\ell)$ be as in corollary 1.7.7 and $D(2)$ be as in the statement of theorem 1.9.1. The following hold:*

1. *$S = \prod_\ell S_\ell$.*
2. *For $\ell \in \mathcal{P}$, $\ell \neq 2$, we have*
$$\left[\mathrm{SL}_2(\mathbb{Z}_\ell) : S_\ell\right] \mid |\mathrm{SL}_2(\mathbb{F}_\ell)| \cdot B(\ell);$$
for $\ell = 2$ we have
$$\left[\mathrm{SL}_2(\mathbb{Z}_2) : S_2\right] < 2^{258} D(2)^{144}.$$
3. *For $\ell \notin \mathcal{P}$ the equality $S_\ell = \mathrm{SL}_2(\mathbb{Z}_\ell)$ holds.*

Proof. (1) This would follow from [123, Théorème 1], but since we do not need the added generality and the proof is quite short we include it here for the reader's convenience.
Regard S as a closed subgroup of $\prod_\ell S_\ell \subseteq \prod_\ell \mathrm{SL}_2(\mathbb{Z}_\ell) = \mathrm{SL}_2(\widehat{\mathbb{Z}})$. For each finite set of primes B, let $p_B \colon S \to S_B = \prod_{\ell \in B} S_\ell$ be the canonical projection. We plan to show that for every such B containing $\mathcal{P}$ we have $p_B(S) = S_B$. Indeed let us consider the case $B = \mathcal{P}$ first. Our choice of F implies that $S_\ell = \rho_\ell(\mathrm{Gal}(\overline{F}/F))$ is a pro-ℓ group for every $\ell \in \mathcal{P}$: the group S_ℓ has trivial reduction modulo ℓ by construction, and therefore S_ℓ admits the usual congruence filtration by the kernels of the reductions modulo ℓ^k for varying k. Now a pro-ℓ group is obviously pro-nilpotent, so $p_B(S)$ is pro-nilpotent as well and therefore it is the product of its pro-Sylow subgroups (which are just the S_ℓ). To treat the general case we recall some terminology from [119]. Following Serre, we say that a finite simple group Σ *occurs* in the profinite group Y if there exist a closed subgroup Y_1 of Y and an open normal subgroup Y_2 of Y_1 such that $\Sigma \cong Y_1/Y_2$. We also write $\mathrm{Occ}(Y)$ for the set of isomorphism classes of finite simple non abelian groups occurring in Y. From [119, IV-25] we read the following description of the sets $\mathrm{Occ}(\mathrm{GL}_2(\mathbb{Z}_p))$:

- $\mathrm{Occ}(\mathrm{GL}_2(\mathbb{Z}_p)) = \emptyset$ for $p = 2, 3$;
- $\mathrm{Occ}(\mathrm{GL}_2(\mathbb{Z}_5)) = \{A_5\}$;
- $\mathrm{Occ}(\mathrm{GL}_2(\mathbb{Z}_p)) = \{\mathrm{PSL}_2(\mathbb{F}_p), A_5\}$ for $p \equiv \pm 1 \pmod 5$, $p > 5$;
- $\mathrm{Occ}(\mathrm{GL}_2(\mathbb{Z}_p)) = \{\mathrm{PSL}_2(\mathbb{F}_p)\}$ for $p \equiv \pm 2 \pmod 5$, $p > 5$.

Let B be a finite set of primes containing $\mathcal{P}$ and satisfying $p_B(S) = S_B$, and fix a prime $\ell_0 \notin B$. We claim that $p_{B\cup\{\ell_0\}}(S) = S_{B\cup\{\ell_0\}}$. Notice first that $\mathrm{PSL}_2(\mathbb{F}_{\ell_0})$ occurs in S_{ℓ_0} and therefore in $p_{B\cup\{\ell_0\}}(S)$; set $N_{\ell_0} = \ker\left(p_{B\cup\{\ell_0\}}(S) \to p_B(S)\right)$. From the exact sequence
$$1 \to N_{\ell_0} \to p_{B\cup\{\ell_0\}}(S) \to p_B(S) \to 1 \tag{1.3}$$
we see that $\mathrm{Occ}\left(p_{B\cup\{\ell_0\}}(S)\right) = \mathrm{Occ}\left(p_B(S)\right) \cup \mathrm{Occ}\left(N_{\ell_0}\right)$. On the other hand, the only finite non-abelian simple groups that can occur in $p_B(S)$ are A_5 and groups of the form $\mathrm{PSL}_2(\mathbb{F}_\ell)$ for $\ell \neq \ell_0$, so

$\mathrm{PSL}_2(\mathbb{F}_{\ell_0})$ does not occur in $p_B(S)$ (notice that $\mathrm{PSL}_2(\mathbb{F}_{\ell_0}) \ncong A_5$ since $\ell_0 \neq 5$), and therefore it must occur in N_{ℓ_0}. Denote by $\overline{N_{\ell_0}}$ the image of N_{ℓ_0} in $\mathrm{SL}_2(\mathbb{F}_{\ell_0})$. The kernel of $N_{\ell_0} \to \mathrm{SL}_2(\mathbb{F}_{\ell_0})$ is a pro-ℓ_0 group, so $\mathrm{Occ}\big(N_{\ell_0}\big)$ equals $\mathrm{Occ}\big(\overline{N_{\ell_0}}\big)$ and therefore $\overline{N_{\ell_0}}$ projects surjectively onto $\mathrm{PSL}_2(\mathbb{F}_{\ell_0})$. Hence we have $\overline{N_{\ell_0}} = \mathrm{SL}_2(\mathbb{F}_{\ell_0})$ by [119, IV-23, Lemma 2], and by lemma 1.3.15 this implies $N_{\ell_0} = \mathrm{SL}_2(\mathbb{Z}_{\ell_0})$: by (1.3) we then have $p_{B\cup\{\ell_0\}}(S) = p_B(S) \times \mathrm{SL}_2(\mathbb{Z}_{\ell_0})$ as claimed. By induction, the equality $p_B(S) = S_B$ holds for any finite set of primes B containing $\mathcal{P}$, and since S is profinite we deduce that $S = \prod_\ell S_\ell$.

(2) The group S_ℓ is the kernel of the projection map $(G_\ell \cap \mathrm{SL}_2(\mathbb{Z}_\ell)) \to \mathrm{SL}_2(\mathbb{F}_\ell)$; as such, it contains the intersection $H'_\ell \cap B_\ell(1)$ (notation as in section 1.7), so we just need to invoke corollary 1.7.7 to have

$$\big[\mathrm{SL}_2(\mathbb{Z}_\ell) : S_\ell\big] \,\big|\, \big[\mathrm{SL}_2(\mathbb{Z}_\ell) : (H'_\ell \cap B_\ell(1))\big] \,\big|\, |\mathrm{SL}_2(\mathbb{F}_\ell)|B(\ell)$$

as claimed. On the other hand, for $\ell = 2$ the group H_2 is a subgroup of $\rho_2(\mathrm{Gal}\left(\overline{K}/K(E[4])\right))$, while S_2 is $\rho_2(\mathrm{Gal}\left(\overline{K}/K_{cyc}(E[2])\right))$, so S_2 is larger than $H'_2 \cap \mathcal{B}_2(1)$ and we can again use the bound of corollary 1.7.7, which now reads

$$\big[\mathrm{SL}_2(\mathbb{Z}_2) : S_2\big] \leq 2^{255} D(2)^{144} |\mathrm{SL}_2(\mathbb{F}_2)| < 2^{258} D(2)^{144}.$$

(3) As $\ell \notin \mathcal{P}$ we know that $\rho_\ell(\mathrm{Gal}(\overline{K}/K))$ contains $\mathrm{SL}_2(\mathbb{Z}_\ell)$, so the group $\mathrm{PSL}_2(\mathbb{F}_\ell)$ occurs in $\rho_\ell(\mathrm{Gal}(\overline{K}/K))$. Consider the Galois group $\mathrm{Gal}(F/K)$: it is by construction a subquotient of $\prod_{p\in\mathcal{P}} \mathrm{GL}_2(\mathbb{Z}_p)$, so the only groups that can occur in it are those in $\bigcup_{p\in\mathcal{P}} \mathrm{Occ}\left(\mathrm{GL}_2(\mathbb{Z}_p)\right)$, and in particular $\mathrm{PSL}_2(\mathbb{F}_\ell)$ does not occur in $\mathrm{Gal}(F/K)$. Now $\rho_\ell(\mathrm{Gal}(\overline{K}/K))$ is an extension of a quotient of $\mathrm{Gal}(F/K)$ by $\rho_\ell\left(\mathrm{Gal}\left(\overline{K}/F\right)\right)$, so $\mathrm{PSL}_2(\mathbb{F}_\ell)$ occurs in $\rho_\ell\left(\mathrm{Gal}\left(\overline{K}/F\right)\right)$, and furthermore $\rho_\ell\left(\mathrm{Gal}\left(\overline{K}/F\right)\right)$ is an extension of an abelian group by $\rho_\ell\left(\mathrm{Gal}\left(\overline{K}/F_{cyc}\right)\right)$, so $\mathrm{PSL}_2(\mathbb{F}_\ell)$ also occurs in $\rho_\ell\left(\mathrm{Gal}\left(\overline{K}/F_{cyc}\right)\right) = S_\ell$: reasoning as in (1), we then see that S_ℓ projects surjectively onto $\mathrm{PSL}_2(\mathbb{F}_\ell)$, and therefore $S_\ell = \mathrm{SL}_2(\mathbb{Z}_\ell)$. □

The proof of theorem 1.9.1 is now immediate:

Proof of theorem 1.9.1. We have already seen that the index $\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right]$ equals $[\mathbb{Z}^\times : \det\circ\rho_\infty \mathrm{Gal}(\overline{K}/K)] \cdot [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)]$. Now the first factor in this product is at most $[K : \mathbb{Q}]$, while the second is bounded by $[\mathrm{SL}_2(\widehat{\mathbb{Z}}) : S]$; it follows that the adelic index is bounded by

$$\begin{aligned}
[K : \mathbb{Q}] \cdot [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : S] &\leq [K : \mathbb{Q}] \cdot \prod_{\ell\in\mathcal{P}} [\mathrm{SL}_2(\mathbb{Z}_\ell) : S_\ell] \\
&\leq [K : \mathbb{Q}] \cdot \prod_{\ell|\Psi} [\mathrm{SL}_2(\mathbb{Z}_\ell) : S_\ell] \\
&< [K : \mathbb{Q}] \cdot 2^{258} \cdot D(2)^{144} \cdot \prod_{\ell|\Psi,\ell\neq 2} |\mathrm{SL}_2(\mathbb{F}_\ell)| \cdot \prod_{\ell|\Psi,\ell\neq 2} B(\ell),
\end{aligned} \tag{1.4}$$

where we have used the fact that $\ell \nmid \Psi \Rightarrow \ell \notin \mathcal{P}$. We now observe that by construction for all odd primes ℓ we have $v_\ell(D(\infty)) \geq v_\ell(D(\ell))$, so by corollary 1.7.7 the quantity $\prod_{\ell|\Psi,\ell\neq 2} B(\ell)$ divides

$$\prod_{\ell|\Psi,\ell\neq 2} \ell^{33}\ell^{48v_\ell(D(\ell))} \,\Big|\, \prod_{\ell|\Psi,\ell\neq 2} \ell^{33}\ell^{48v_\ell(D(\infty))},$$

which in turn divides $\left(\frac{\mathrm{rad}(\Psi)}{2}\right)^{33} \cdot D(\infty)^{48}$. Combining this fact with equation (1.4) and the trivial bound $|\mathrm{SL}_2(\mathbb{F}_\ell)| < \ell^3$ we find that the adelic index is at most

$$[K:\mathbb{Q}] \cdot 2^{225} \cdot D(2)^{144} \cdot \left(\prod_{\ell|\Psi, \ell\neq 2} \ell^3 \right) \cdot \mathrm{rad}(\Psi)^{33} \cdot D(\infty)^{48},$$

which in turn is less than $[K:\mathbb{Q}] \cdot 2^{222} \cdot D(2)^{144} \cdot \mathrm{rad}(\Psi)^{36} \cdot D(\infty)^{48}$, whence the theorem. □

Using the estimates of proposition 1.2.6 to bound $\Psi, D(2)$ and $D(\infty)$ we get:

Corollary 1.9.3. *(Theorem 1.1.1) Let E/K be an elliptic curve that does not admit complex multiplication. The inequality*

$$\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right] < \gamma_1 \cdot [K:\mathbb{Q}]^{\gamma_2} \cdot \max\{1, h(E), \log[K:\mathbb{Q}]\}^{2\gamma_2}$$

holds, where $\gamma_1 = \exp(10^{21483})$ and $\gamma_2 = 2.4 \cdot 10^{10}$.

Remark 1.9.4. With some work, the techniques used in [62] (cf. especially Theorem 4.2 of *op. cit.*) could be used to improve the above bound on Ψ; unfortunately, the same methods do not seem to be easily applicable to bound $D(\infty)$. Notice that our estimates for Ψ and $D(\infty)$ are essentially of the same order of magnitude, so using a finer bound for Ψ without changing the one for $D(\infty)$ would only yield a minor improvement of the final result.
On the other hand, it is easy to see that using the improved version of the isogeny theorem mentioned in remarks 1.2.3 and 1.2.7 one can prove

$$\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\right] < \gamma_3 \cdot ([K:\mathbb{Q}] \cdot \max\{1, h(E), \log[K:\mathbb{Q}]\})^{\gamma_4}$$

with $\gamma_3 = \exp\left(1.9 \cdot 10^{10}\right)$ and $\gamma_4 = 12395$.

1.9.1 The field generated by a torsion point

As an easy consequence of our main result we can also prove:

Corollary 1.1.3. *Let E/K be an elliptic curve that does not admit complex multiplication. There exists a constant $\gamma(E/K)$ with the following property: for every $x \in E_{\mathrm{tors}}(\overline{K})$ (of order denoted $N(x)$) the inequality*

$$[K(x):K] \geq \gamma(E/K) N(x)^2$$

holds. We can take $\gamma(E/K) = \left(\zeta(2) \cdot \left[\,\mathrm{GL}_2\left(\widehat{\mathbb{Z}}\right) : \rho_\infty\left(\mathrm{Gal}(\overline{K}/K)\right)\,\right]\right)^{-1}$, which can be explicitly bounded thanks to the main theorem.

Proof. For any such x set $N = N(x)$ and choose a point $y \in E[N]$ such that (x, y) is a basis of $E[N]$ as $(\mathbb{Z}/N\mathbb{Z})$-module. Let $G(N)$ be the image of $\mathrm{Gal}(\overline{K}/K)$ inside $\mathrm{Aut}\, E[N]$, which we identify with $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ via the basis (x, y). We have a tower of extensions $K(E[N])/K(x)/K$, where $K(E[N])$ is Galois over K and therefore over $K(x)$. The Galois groups of these extensions are given – essentially by definition – by

$$\mathrm{Gal}(K(E[N])/K) = G(N), \quad \mathrm{Gal}(K(E[N])/K(x)) = \mathrm{Stab}(x),$$

where $\mathrm{Stab}(x) = \{\sigma \in G(N) \mid \sigma(x) = x\}$. It follows that

$$[K(x) : K] = \frac{[K(E[N]) : K]}{[K(E[N]) : K(x)]} = \frac{|G(N)|}{|\mathrm{Stab}(x)|},$$

and furthermore it is easy to check that

$$|G(N)| = \frac{|\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})|}{[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : G(N)]} = \frac{N^3\varphi(N)\prod_{p|N}\left(1 - \frac{1}{p^2}\right)}{[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : G(N)]}.$$

On the other hand, the stabilizer of x in $G(N)$ is contained in the stabilizer of x in $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, which is simply

$$\left\{ \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \mid a \in \mathbb{Z}/N\mathbb{Z},\ b \in (\mathbb{Z}/N\mathbb{Z})^\times \right\},$$

so $|\mathrm{Stab}(x)| \leq |\mathbb{Z}/N\mathbb{Z}| \cdot \left|(\mathbb{Z}/N\mathbb{Z})^\times\right| = N\varphi(N)$. Finally, the index of $G(N)$ inside $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ is certainly not larger than the index of G_∞ inside $\mathrm{GL}_2(\widehat{\mathbb{Z}})$. Putting everything together we obtain

$$[K(x) : K] = \frac{N^3\varphi(N)\prod_{p|N}\left(1 - \frac{1}{p^2}\right)}{[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : G(N)] \cdot |\mathrm{Stab}(x)|} \geq \frac{N^3\varphi(N)\prod_{p \text{ prime}}\left(1 - \frac{1}{p^2}\right)}{N\varphi(N) \cdot [\mathrm{GL}_2(\widehat{\mathbb{Z}}) : G_\infty]},$$

and the corollary follows by remarking that $\prod_{p \text{ prime}}\left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)}$. □

Chapter 2

Products of Elliptic Curves

2.1 Introduction

In this work we prove an explicit, adelic surjectivity result for the Galois representation attached to a product of pairwise non-isogenous, non-CM elliptic curves, extending the result of chapter 1. Our main theorem is as follows:

Theorem 2.1.1. *Let $E_1, \ldots, E_n$, $n \geq 2$, be elliptic curves defined over a number field K, pairwise not isogenous over $\overline{K}$. Suppose that $\mathrm{End}_{\overline{K}}(E_i) = \mathbb{Z}$ for $i = 1, \ldots, n$, and denote by G_∞ the image of $\mathrm{Gal}(\overline{K}/K)$ inside*

$$\prod_{i=1}^{n} \prod_{\ell} \mathrm{Aut}(T_\ell(E_i)) \subset \mathrm{GL}_2(\hat{\mathbb{Z}})^n.$$

Set $\gamma := 10^{13}, \delta := \exp\exp\exp(13)$, and let $H = \max\{1, \log[K:\mathbb{Q}], \max_i h(E_i)\}$, where $h(E_i)$ denotes the stable Faltings height of E_i. The group G_∞ has index at most

$$\delta^{n(n-1)} \cdot \left([K:\mathbb{Q}] \cdot H^2\right)^{\gamma n(n-1)}$$

in

$$\Delta := \left\{ (x_1, \ldots, x_n) \in \mathrm{GL}_2(\hat{\mathbb{Z}})^n \;\middle|\; \det x_i = \det x_j \;\; \forall i, j \right\}.$$

Remark 2.1.2. Note that the compatibility of the Weil pairing with the action of Galois forces G_∞ to be contained in Δ. Also note that the statements we actually prove (lemma 2.7.3 and theorem 2.7.5 below) are slightly more precise, and immediately imply theorem 2.1.1 by proposition 2.2.5 and elementary estimates.

It should be noted that it has been known since the work of Serre and Masser-Wüstholz (cf. [71], Main Theorem and Proposition 1) that the isogeny theorem (section 2.2 below) gives an effective bound ℓ_0 on the largest prime ℓ for which the image of the representation

$$\mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(T_\ell(E_1 \times \ldots \times E_n))$$

does not contain $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$. As it was in chapter 1, the main difficulty here lies in controlling the image of the representation modulo powers of primes smaller than ℓ_0.

The proof of theorem 2.1.1 is somewhat technical, so before fiddling with the details we describe the main ideas behind it. The general framework is the same as that of the proof of the non-effective open image theorem for such a product (cf. for example [108, Theorem 3.5]), with the

added difficulties that naturally arise when trying to actually compute the index. In particular, when writing 'of finite index' or 'open' in the sketch that follows we tacitly imply that the index in question is explicitly computable in terms of the data. In those instances when the need will arise to actually quantify indices, it will be useful to work with the following 'standard' open subgroups:

Definition 2.1.3. For a prime ℓ and a positive integer s we let $\mathcal{B}_\ell(s)$ be the open subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ given by

$$\left\{x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; x \equiv \mathrm{Id} \pmod{\ell^s}\right\}.$$

We also set $\mathcal{B}_\ell(0) = \mathrm{SL}_2(\mathbb{Z}_\ell)$, and for non-negative integers $k_1, \ldots, k_n$ we denote by $\mathcal{B}_\ell(k_1, \ldots, k_n)$ the open subgroup $\prod_{j=1}^n \mathcal{B}_\ell(k_i)$ of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$.

Let us now describe the proof method proper.

It is not hard to see that it is enough to consider the intersection $G_\infty \cap \mathrm{SL}_2(\mathbb{Z}_\ell)^n$, because the determinant of G_∞ agrees with the cyclotomic character and is therefore well understood. A short argument then shows that it is enough to consider products $E_1 \times E_2$ involving only two factors: this is done by proving that a subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$ whose projection on any *pair* of factors is of finite index is itself of finite (and explicitly bounded) index. This step will be carried out in section 2.3 below, and should be thought of as the 'integral' version of [109, Lemma on p. 790].

With this result at hand we are thus reduced to dealing with subgroups G of $\mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell)$ whose projections on either factor are of finite index in $\mathrm{SL}_2(\mathbb{Z}_\ell)$. Note that the fact that this index is finite is the open image theorem for a single elliptic curve, which was proved by Serre in [116] and made explicit in chapter 1. We wish to show that G is of (explicitly bounded) finite index in $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$, that is, we want to exhibit a t such that G contains $\mathcal{B}_\ell(t,t)$: this clearly comes down to proving that the two kernels $K_i = \ker\left(G \xrightarrow{\pi_i} \mathrm{SL}_2(\mathbb{Z}_\ell)\right)$, when identified with subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)$, are of (explicitly bounded) finite index. By symmetry, we just need to deal with K_1.

In section 2.4 we linearize the problem by reducing it to the study of certain $\mathbb{Z}_\ell$-Lie algebras. We also give the statements of two technical results whose proof, being rather lengthy, is deferred to chapter 8; while the results themselves are more complicated, the methods used to show them do not differ much from those of chapter 1, where the case of a single elliptic curve is treated.

A simple lemma, again given in section 2.4, further reduces the problem of finding an integer t such that $\mathcal{B}_\ell(t)$ is contained in K_1 to the (easier) question of finding a t such that $K_1(\ell^t)$, the reduction modulo ℓ^t of K_1, is nontrivial. We exploit here the fact that $\pi_2(G)$ (the projection of G on the second factor $\mathrm{SL}_2(\mathbb{Z}_\ell)$) acts by conjugation on K_1, the latter being a normal subgroup of G: we prove that a group whose reduction modulo ℓ^t is nontrivial and that is stable under conjugation by a finite-index subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ must itself be of finite index in $\mathrm{SL}_2(\mathbb{Z}_\ell)$. This reduction step is made simpler by the fact that we can work with Lie algebras instead of treating directly the corresponding groups (which might be quite complicated).

Next we ask what happens if we suppose that the smallest integer t such that $K_1(\ell^t)$ is nontrivial is in fact very large. The conclusion is that the Lie algebra of G looks 'very much like' the graph of a Lie algebra morphism $\mathfrak{sl}_2(\mathbb{Z}_\ell) \to \mathfrak{sl}_2(\mathbb{Z}_\ell)$, namely it induces an actual Lie algebra morphism when regarded modulo ℓ^N for a very large N (depending on t). Following for example the approach of Ribet (cf. the theorems on p. 795 of [109]), we would like to know that all such morphisms are

'inner', that is, they are given by conjugation by a certain matrix: it turns out that this is also true in our context, even though the result is a little more complicated to state (cf. section 2.5).
In section 2.6 we then deal with the case of two elliptic curves, applying the aforementioned results to deduce an open image theorem for each prime ℓ. It is then an easy matter to deduce, in section 2.7, the desired adelic result for any finite product.

Notation. Throughout the whole chapter, the prime 2 plays a rather special role, and special care is needed to treat it. In order to give uniform statements that hold for every prime we put $v = 0$ or 1 according to whether the prime ℓ we are working with is odd or equals 2, that is we set

$$v = v_\ell(2) = \begin{cases} 0, & \text{if } \ell \text{ is odd} \\ 1, & \text{otherwise.} \end{cases}$$

We will also consistently use the following notations:

- G_ℓ, to denote the image of $\mathrm{Gal}(\overline{K}/K)$ in $\mathrm{Aut}\, T_\ell(E_1) \times \cdots \times \mathrm{Aut}\, T_\ell(E_k)$;
- $G(\ell^n)$, where G is a closed subgroup of a certain $\mathrm{GL}_2(\mathbb{Z}_\ell)^k$, to denote the reduction of G modulo ℓ^n, that is to say its image in $\mathrm{GL}_2(\mathbb{Z}/\ell^n\mathbb{Z})^k$;
- $N(G)$, to denote the largest normal pro-ℓ subgroup of G;
- G', to denote the topological closure of the commutator subgroup.

2.2 Preliminaries on isogeny bounds

The main tool that makes all the effective estimates possible is the isogeny theorem of Masser and Wüstholz [70] [72], which we employ in the explicit version proved in [28]. We need some notation: we let $\alpha(g) = 2^{10}g^3$ and define, for any abelian variety A/K of dimension g,

$$b(A/K) = b([K:\mathbb{Q}], g, h(A)) = \left((14g)^{64g^2}[K:\mathbb{Q}]\max\left(h(A), \log[K:\mathbb{Q}], 1\right)^2\right)^{\alpha(g)}.$$

Theorem 2.2.1. *([28, Théorème 1.4]) Let K be a number field and A, A^* be two Abelian K-varieties of dimension g. If A, A^* are isogenous over K, then there exists a K-isogeny $A^* \to A$ whose degree is bounded by $b([K:\mathbb{Q}], \dim(A), h(A))$.*

Remark 2.2.2. As the notation suggests, the three arguments of b will always be the degree of a number field K, the dimension g of an Abelian variety A/K and its stable Faltings height $h(A)$.

In [68] (cf. especially lemma 3.4) Masser shows the following:

Theorem 2.2.3. *(Masser) Suppose that A/K is an Abelian variety that is isomorphic over $\overline{K}$ to a product $A_1^{e_1} \times \ldots \times A_m^{e_m}$, where each A_i is simple and has trivial endomorphism ring over $\overline{K}$. Suppose furthermore that for every A^* isogenous to A over K we can find an isogeny $A^* \to A$ of degree bounded by b for a certain constant b. Then there exists an integer $b_0 \leq b$ such that we can always choose a K-isogeny $A^* \to A$ of degree dividing b_0.*

We will denote by $b_0(A/K)$ the minimal b_0 with the property of the above theorem; in particular $b_0(A/K) \leq b(A/K)$. Consider now $b_0(A/K')$ as K' ranges through all the finite extensions of K of degree bounded by d. On one hand, $b_0(A/K)$ divides $b_0(A/K')$ ([68], p.190); on the other $b_0(A/K') \leq b(d[K:\mathbb{Q}], h(A), \dim(A))$ stays bounded, and therefore the number

$$\mathrm{lcm}_{[K':K]\leq d}\, b_0(A/K')$$

exists and is finite. We give this function a name:

Definition 2.2.4. Suppose A/K is a product of simple varieties with absolutely trivial endomorphism ring. Then we define

$$b_0(A/K;d) = \mathrm{lcm}_{[K':K]\leq d}\, b_0(A/K').$$

The function $b_0(A/K;d)$ is studied in [68, Theorem D]. Adapting the argument given by Masser to the form of the function $b(d[K:\mathbb{Q}], h(A), \dim(A))$ at our disposal it is immediate to prove:

Proposition 2.2.5. *(proposition 1.2.6) If A/K is as in the previous definition and of dimension g, then*

$$b_0(A/K;d) \leq b(A/K;d) := 4^{\exp(1)\cdot(d(1+\log d)^2)^{\alpha(g)}} b([K:\mathbb{Q}], \dim(A), h(A))^{1+\alpha(g)\log(d(1+\log d)^2)}.$$

2.3 An integral Goursat-Ribet lemma for $\mathrm{SL}_2(\mathbb{Z}_\ell)$

As anticipated, we show that a (necessary and) sufficient condition for a subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$ to be open is that all its projections on pairs of factors $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$ are themselves open. This will follow rather easily from the following elementary lemma (whose easy verification we omit):

Lemma 2.3.1. *Let s_1, s_2 be non-negative integers (with $s_1, s_2 \geq 2$ if $\ell = 2$ and $s_1, s_2 \geq 1$ if $\ell = 3$). The commutator group $[\mathcal{B}_\ell(s_1), \mathcal{B}_\ell(s_2)]$ contains $B_\ell(s_1+s_2+v)$, and the iterated commutator $\underbrace{[\cdots[}_{(n-1)\text{ times}} \mathcal{B}_\ell(s_1), \mathcal{B}_\ell(s_2)\,], \mathcal{B}_\ell(s_3)], \cdots, \mathcal{B}_\ell(s_n)]$ contains $B_\ell(s_1 + \cdots + s_n + (n-1)v)$.*

Lemma 2.3.2. *Let n be a positive integer, G a closed subgroup of $\prod_{i=1}^n \mathrm{SL}_2(\mathbb{Z}_\ell)$, and π_i the projection from G on the i-th factor. Suppose that, for every $i \neq j$, the group $(\pi_i \times \pi_j)(G)$ contains $\mathcal{B}_\ell(s_{ij}, s_{ij})$ for a certain non-negative integer s_{ij} (with $s_{ij} \geq 2$ if $\ell = 2$ and $s_{ij} \geq 1$ if $\ell = 3$): then G contains $\prod_{i=1}^n \mathcal{B}_\ell\left(\sum_{j\neq i} s_{ij} + (n-1)v\right)$.*

Proof. Clearly by the symmetry of the problem it is enough to show that G contains

$$\{\mathrm{Id}\} \times \cdots \times \{\mathrm{Id}\} \times \mathcal{B}_\ell\left(\sum_{j\neq n} s_{nj} + (n-1)v\right).$$

By lemma 2.3.1, for any g in $\mathcal{B}_\ell\left(\sum_{j\neq n} s_{nj} + (n-1)v\right)$ there exist elements y_i in $\mathcal{B}_\ell(s_{ni})$ (for i between 1 and $n-1$) such that g can be written as $[\cdots[[y_1, y_2], y_3], \cdots, y_{n-1}]$. By hypothesis we can find $x_1, \ldots, x_{n-1} \in G$ such that $\pi_i(x_i) = \mathrm{Id}$ and $\pi_n(x_i) = y_i$ for all i between 1 and $n-1$. Consider now the iterated commutator

$$\tilde{g} = [\cdots[[x_1, x_2], x_3], \cdots, x_{n-1}] :$$

this is a product of elements of G, and therefore it is itself an element of G. For $i \leq n-1$, the i-th component of $\tilde{g}$ is trivial, since

$$\begin{aligned}\pi_i(\tilde{g}) &= [\cdots[\cdots[[\pi_i(x_1),\pi_i(x_2)],\pi_i(x_3)],\cdots,\underbrace{\pi_i(x_i)}_{\mathrm{Id}}],\cdots,\pi_i(x_{n-1})] \\ &= [\cdots[\cdots[[\pi_i(x_1),\pi_i(x_2)],\pi_i(x_3)],\cdots,\mathrm{Id}],\cdots,\pi_i(x_{n-1})] \\ &= \mathrm{Id}\,.\end{aligned}$$

On the other hand, our choice of $y_1,\ldots,y_{n-1}$ ensures that $\pi_n(\tilde{g}) = [\cdots[[y_1,y_2],y_3],\cdots,y_{n-1}] = g$. We have thus shown that $(1,1,\ldots,1,g) = \tilde{g}$ is an element of G for any choice of g in

$$\mathcal{B}_\ell\left(\sum_{j\neq n} s_{ij} + (n-1)v\right),$$

and repeating the argument for the other projections gives the required result. □

Corollary 2.3.3. *Let G be a closed subgroup of $\prod_{i=1}^n \mathrm{SL}_2(\hat{\mathbb{Z}})$ with $n \geq 2$. For every pair of indices $i \neq j$ let $S^{(i,j)}$ be a subgroup of $\mathrm{SL}_2(\hat{\mathbb{Z}})^2$ with the following properties:*

- *the projection of G on the direct factor $\mathrm{SL}_2(\hat{\mathbb{Z}}) \times \mathrm{SL}_2(\hat{\mathbb{Z}})$ corresponding to the pair of indices (i,j) contains $S^{(i,j)}$;*
- *$S^{(i,j)}$ decomposes as a direct product $\prod_{\ell\ prime} S_\ell^{(i,j)} \subseteq \prod_\ell \mathrm{SL}_2(\mathbb{Z}_\ell)^2$;*
- *for almost every ℓ, the group $S_\ell^{(i,j)}$ is all of $\mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell)$;*
- *for every prime ℓ such that $S_\ell^{(i,j)} \neq \mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell)$ there exists an integer $f_\ell^{(i,j)}$ such that $S_\ell^{(i,j)} = \mathcal{B}_\ell(f_\ell^{(i,j)}, f_\ell^{(i,j)})$ (if $\ell = 2$ we demand that $f_2^{(i,j)} \geq 2$, while if $\ell = 3$ we impose the condition $f_3^{(i,j)} \geq 1$).*

Denote by $c^{(i,j)}$ the index of $S^{(i,j)}$ in $\mathrm{SL}_2(\hat{\mathbb{Z}}) \times \mathrm{SL}_2(\hat{\mathbb{Z}})$ and $c = \max_{i\neq j} c^{(i,j)}$. The index of G in $\prod_{i=1}^n \mathrm{SL}_2(\hat{\mathbb{Z}})$ is strictly less than

$$(8\zeta(2))^{n(n-1)} c^{n(n-1)/2}.$$

Proof. Let $\ell > 3$ be a prime. If $S_\ell^{(i,j)} = \mathrm{SL}_2(\mathbb{Z}_\ell)^2$ for all (i,j), then the previous lemma applies (with $s_{ij} = 0$ for every pair of indices (i,j)) and shows that $\prod_{k=1}^n \mathrm{SL}_2(\mathbb{Z}_\ell)$ is contained in G. Suppose on the other hand that either $\ell \leq 3$ or for at least one pair (i,j) we have $S_\ell^{(i,j)} \neq \mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell)$. The previous lemma tells us that the projection of G on the direct factor $\prod_{i=1}^n \mathrm{SL}_2(\mathbb{Z}_\ell)$ of $\prod_{i=1}^n \mathrm{SL}_2(\hat{\mathbb{Z}})$ contains

$$B_\ell\left(\sum_{j\neq 1} f_\ell^{(1,j)} + (n-1)v, \ldots, \sum_{j\neq n} f_\ell^{(n,j)} + (n-1)v\right).$$

Notice that the index of this group in $\prod_{i=1}^n \mathrm{SL}_2(\mathbb{Z}_\ell)$ is at most

$$\prod_{i=1}^n \left(\ell^{3\sum_{j\neq i} f_\ell^{(i,j)} + 3(n-1)v}\right) = 2^{3n(n-1)v} \prod_{i=1}^n \prod_{j\neq i} \ell^{3f_\ell^{(i,j)}}.$$

Let now $\mathcal{P} = \{2,3\} \cup \left\{ \ell \;\middle|\; \exists (i,j) : S_\ell^{(i,j)} \neq \mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell) \right\}$. By what we have just seen,

$$\left[\prod_{k=1}^{n} \mathrm{SL}_2(\hat{\mathbb{Z}}) : G \right] \leq 2^{3n(n-1)} \prod_{\ell \in \mathcal{P}} \prod_{i=1}^{n} \prod_{j \neq i} \ell^{3 f_\ell^{(i,j)}}.$$

On the other hand, note that the index of $S_\ell^{(i,j)}$ in $\mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell)$ is at least $\ell^{6 f_\ell^{(i,j)}} \cdot \left(\frac{\ell^2-1}{\ell^2} \right)^2$, so the above product is bounded by

$$\begin{aligned}
2^{3n(n-1)} \prod_{\ell \in \mathcal{P}} \prod_{i<j} & \left\{ \left[\mathrm{SL}_2(\mathbb{Z}_\ell)^2 : S_\ell^{(i,j)} \right] \cdot \left(\frac{\ell^2}{\ell^2-1} \right)^2 \right\} \\
& \leq 2^{3n(n-1)} \prod_{\ell} \left(\frac{\ell^2}{\ell^2-1} \right)^{n(n-1)} \cdot \prod_{i<j} \prod_{\ell \in \mathcal{P}} \left[\mathrm{SL}_2(\mathbb{Z}_\ell)^2 : S_\ell^{(i,j)} \right] \\
& \leq 2^{3n(n-1)} \zeta(2)^{n(n-1)} \prod_{i<j} c^{(i,j)} \\
& \leq 2^{3n(n-1)} \zeta(2)^{n(n-1)} c^{n(n-1)/2}.
\end{aligned}$$

□

2.4 Lie subalgebras of $\mathrm{sl}_2(\mathbb{Z}_\ell)^n$ and some Pink-type results

Let us briefly recall the construction (essentially due to Pink) of the $\mathbb{Z}_\ell$-Lie algebra associated with a subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$:

Definition 2.4.1. (cf. [97]) Let ℓ be a prime. Define maps Θ_n as follows:

$$\begin{array}{rccc} \Theta_n : & \mathrm{GL}_2(\mathbb{Z}_\ell)^n & \to & \bigoplus_{i=1}^n \mathfrak{sl}_2(\mathbb{Z}_\ell) \\ & (g_1, \ldots, g_n) & \mapsto & \left(g_1 - \frac{1}{2}\operatorname{tr}(g_1), \ldots, g_n - \frac{1}{2}\operatorname{tr}(g_n) \right). \end{array}$$

If G is a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$ (resp. of $B_2(2,\ldots,2)$ in case $\ell = 2$), define $L(G) \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)^n$ to be the $\mathbb{Z}_\ell$-span of $\Theta_n(G)$. We call $L(G)$ the **Lie algebra** of G.

The crucial importance of this construction lies in the fact that it allows us to linearize the problem of showing that a certain subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$ contains an open neighbourhood of the identity: indeed, we have the following two results, for whose proof we refer the reader to chapter 8.

Theorem 2.4.2. *Let $\ell > 2$ be a prime number and G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell) \times \mathrm{GL}_2(\mathbb{Z}_\ell)$. Let G_1, G_2 be the two projections of G on the two factors $\mathrm{GL}_2(\mathbb{Z}_\ell)$, and let n_1, n_2 be positive integers such that G_i contains $\mathcal{B}_\ell(n_i)$ for $i = 1, 2$. Suppose furthermore that for every $(g_1, g_2) \in G$ we have $\det(g_1) = \det(g_2)$. At least one of the following holds:*

- *G contains $\mathcal{B}_\ell(20\max\{n_1,n_2\}, 20\max\{n_1,n_2\})$*
- *there exists a subgroup T of G, of index dividing $2 \cdot 48^2$, with the following properties:*
  - *if $L(T)$ contains $\ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell)$ for a certain integer k, then T contains $\mathcal{B}_\ell(p,p)$, where*

$$p = 2k + \max\{2k+4, 8n_1, 8n_2\}.$$

  We call this property $(*)$.

– *for any (t_1, t_2) in T, if both $[t_1]$ and $[t_2]$ are multiples of the identity, then they are equal.*

Theorem 2.4.3. *Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_2) \times \mathrm{GL}_2(\mathbb{Z}_2)$ whose projection modulo 4 is trivial. Denote by G_1, G_2 the two projections of G on the factors $\mathrm{GL}_2(\mathbb{Z}_2)$, and let $n_1 \geq 4, n_2 \geq 4$ be integers such that G_i contains $\mathcal{B}_2(n_i)$. Suppose furthermore that for every $(g_1, g_2) \in G$ we have $\det(g_1) = \det(g_2)$. If $L(G)$ contains $2^k\mathfrak{sl}_2(\mathbb{Z}_2) \oplus 2^k\mathfrak{sl}_2(\mathbb{Z}_2)$ for a certain $k \geq 2$, then G contains*

$$\mathcal{B}_2(12(k + 12n_2 + 5n_1 + 13) + 1, 12(k + 12n_1 + 5n_2 + 13) + 1).$$

Finally, the following easy lemma characterizes conjugation-stable subalgebras of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$:

Lemma 2.4.4. *(Lemma 8.2.1) Let ℓ be a prime number, t a non-negative integer, and $W \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)$ a Lie subalgebra that does not reduce to zero modulo ℓ^{t+1} and that is stable under conjugation by $\mathcal{B}_\ell(s)$, where $s \geq 0$ is at least 2 if $\ell = 2$ and at least 1 if $\ell = 3$ or 5 (no conditions are necessary if $\ell \geq 7$). The open set $\ell^{t+4s+4v}\mathfrak{sl}_2(\mathbb{Z}_\ell)$ is contained in W.*

2.5 The automorphisms of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ are inner

We will obtain in this section a description of the automorphisms of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ showing that they are all inner, in a suitable sense. In order to establish the required result we first need a few simple preliminaries, starting with the following well-known version of Hensel's lemma:

Lemma 2.5.1. *Let $p(x) \in \mathbb{Z}_\ell[x]$ be a monic polynomial and let α be an element of $\mathbb{Z}_\ell$. Suppose that $v_\ell(p(\alpha)) > 2v_\ell(p'(\alpha))$: then $p(x)$ admits a root $\bar{\alpha}$ such that $v_\ell(\alpha - \bar{\alpha}) \geq v_\ell(p(\alpha)) - v_\ell(p'(\alpha))$.*

The main tool we will use to produce approximate roots of polynomials is the following lemma:

Lemma 2.5.2. *Let ℓ be a prime number, $n \geq 1, m \geq 1$, $g \in \mathrm{End}\,(\mathbb{Z}_\ell^m)$ and $p_g(t)$ the characteristic polynomial of g. Let furthermore $\lambda \in \mathbb{Z}_\ell, w \in \mathbb{Z}_\ell^m$ be such that $gw \equiv \lambda w \pmod{\ell^n}$. Suppose that at least one of the coordinates of w has ℓ-adic valuation at most α: then $p_g(\lambda) \equiv 0 \pmod{\ell^{n-\alpha}}$.*

Proof. Denote by $(g - \lambda\,\mathrm{Id})^*$ the adjugate matrix of $(g - \lambda\,\mathrm{Id})$, that is the unique operator such that $(g - \lambda\,\mathrm{Id})^*(g - \lambda\,\mathrm{Id}) = \det(g - \lambda\,\mathrm{Id}) \cdot \mathrm{Id}$. Multiplying $(g - \lambda\,\mathrm{Id})w \equiv 0 \pmod{\ell^n}$ on the left by $(g - \lambda\,\mathrm{Id})^*$ we obtain $\det(g - \lambda\,\mathrm{Id}) \cdot \mathrm{Id}\, w \equiv 0 \pmod{\ell^n}$, and by considering the coordinate of w of smallest valuation we have $p_g(\lambda) = \det(g - \lambda\,\mathrm{Id}) \equiv 0 \pmod{\ell^{n-\alpha}}$ as claimed. □

An immediate computation also shows:

Lemma 2.5.3. *Let $g \in \mathfrak{sl}_2(\mathbb{Z}_\ell)$. The linear operator $\mathcal{C}_g := [g, \cdot]$ from $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ to itself has eigenvalues $0, \pm 2\mu$, where $\pm\mu$ are the eigenvalues of g, so $p_{\mathcal{C}_g}(t) = t(t^2 - 4\mu^2)$.*

Combining the previous results we obtain the following lemma, which will be very useful for our purposes:

Lemma 2.5.4. *Let g be an element of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$, w be a vector in $\mathbb{Z}_\ell^2$, and β be the minimal valuation of the coefficients of w. Suppose $gw \equiv \lambda w \pmod{\ell^n}$. Then either g has an eigenvalue ν such that $v_\ell(\nu - \lambda) \geq v_\ell(\lambda) + 3$ or else β is at least $n - 2(2 + v_\ell(\lambda))$.*

Proof. Let $\pm\mu$ be the eigenvalues of g. From lemma 2.5.2 we deduce that $v_\ell(p_g(\lambda)) \geq n - \beta$; notice further that $p_g(t) = t^2 - \mu^2$, so $p_g'(t) = 2t$. Suppose that $\beta < n - 2(2 + v_\ell(\lambda))$: then $n - \beta > 2(2 + v_\ell(\lambda)) > 2v_\ell(p_g'(\lambda))$, and by Hensel's lemma $p_g(t)$ has a root ν such that

$$v_\ell(\nu - \lambda) \geq v_\ell(p_g(\lambda)) - v_\ell(p_g'(\lambda)) \geq n - \beta - v - v_\ell(\lambda) \geq v_\ell(\lambda) + 3.$$

□

We now come to the central result of this section, which as anticipated is essentially a description of the Lie algebra automorphisms of (the finite quotients of) $\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

Notation. For the remainder of this section, in order to make notation lighter, when a is a positive integer we write $x = y + O(a)$ for $x \equiv y \pmod{\ell^a}$.

Proposition 2.5.5. *Let L_1 be a subalgebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ and $n \geq 1, s \geq 0$ be integers. Suppose that L_1 contains $\ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell)$ and that $\varphi : L_1 \to \mathfrak{sl}_2(\mathbb{Z}_\ell)$ is a linear map such that*

$$(*) \quad [\varphi(a), \varphi(b)] \equiv \varphi([a, b]) \pmod{\ell^n} \quad \forall a, b \in \ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell).$$

Define

$$x = \varphi\left(\ell^s \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right), \; y = \varphi\left(\ell^s \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right), \; h = \varphi\left(\ell^s \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right)$$

and let α be the minimal integer such that x, y are both nonzero modulo $\ell^{\alpha+1}$.
Suppose that $n \geq \alpha + 10s + 5v + 6$. There exists a matrix $M \in M_2(\mathbb{Z}_\ell)$, at least one of whose coefficients is nonzero modulo ℓ, and such that for every $w \in (\mathbb{Z}_\ell)^2$ and every $g_1 \in L_1$ we have

$$M(g_1 \cdot w) \equiv \varphi(g_1) \cdot M(w) \pmod{\ell^{n-\alpha-6s-4v-6}}. \tag{2.1}$$

Furthermore, $\det(M)$ *does not vanish modulo ℓ^{4s+v}, and for every g_1 in L_1 we have*

$$\operatorname{tr}\left(\varphi(g_1)^2\right) \equiv \operatorname{tr}\left(g_1^2\right) \pmod{\ell^{n-\alpha-10s-5v-6}}$$

and

$$\varphi(g_1) \equiv Mg_1M^{-1} \pmod{\ell^{n-\alpha-10s-5v-6}}, \quad M^{-1}\varphi(g_1)M \equiv g_1 \pmod{\ell^{n-\alpha-10s-5v-6}}$$

Remark 2.5.6. The reader might wonder whether it is really necessary for the three parameters n, α and s to all appear in equation (2.1). The answer is yes. This is apparent for n, if the result is to say something nontrivial about φ. Consider next the limiting case where $\varphi \equiv 0$ (i.e. α goes to infinity): this map satisfies the hypotheses in the proposition for every n, but it is easy to realize that (independently of n) the equality

$$M(g_1 \cdot w) \equiv \varphi(g_1) \cdot M(w) = 0 \pmod{\ell^N}$$

can only hold for bounded N; of course a similar conclusion holds if α stays finite, but is very large. Finally, choose an n and any linear map φ and suppose s is sent to infinity. For s large enough, the condition in the proposition will become void, since both sides of the equality will automatically be 0 modulo ℓ^n: but then we cannot hope to deduce anything meaningful about φ, so that s, too, has to appear in the conclusion.
The question of whether the *dependence* on the parameters is optimal, on the other hand, is far more complicated, and there is almost certainly room for improvement.

Here again let us say a few words about the method of proof before starting with the technical details. To simplify matters, consider the algebra $L = \mathfrak{sl}_2(\mathbb{Q}_\ell)$. Proving that every automorphism of L is inner basically boils down to showing that the only 2-dimensional representation of $\mathfrak{sl}_2(\mathbb{Q}_\ell)$ is the standard one, a result which is usually proved through the 'highest weight vector' machinery: one shows that it is possible to choose an eigenvector v for h that is killed by x, and then describes its full orbit under the action of x, y, h. More precisely, one shows that yv is an eigenvector for h, that xyv is proportional to v, and that $y^2v = 0$.
The proof that follows mimics this very argument by producing a vector v_+, by definition an eigenvector for h, which plays the role of the highest weight vector, and subsequently finding its orbit under the action of h, x, y. The main difficulty lies probably in the initial step, where we need to prove that the eigenvalues of h lie in $\mathbb{Z}_\ell$ and are of a certain shape. Once this is done, most of the proof looks very much like the one for $\mathfrak{sl}_2(\mathbb{Q}_\ell)$, with the additional complication that we have to keep track of valuations along the way.

Proof. Denote by $\mathcal{C}_h$ the linear endomorphism of $\mathfrak{sl}_2(\mathbb{Z}_\ell) \cong \mathbb{Z}_\ell^3$ given by taking the commutator with h. It is clear that

$$\mathcal{C}_h(x) = [h, x] \equiv \varphi\left[\ell^s \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ell^s \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right] \equiv \varphi\left(2\ell^s \cdot \ell^s \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) \equiv 2\ell^s x \pmod{\ell^n},$$

so x is an (approximate) eigenvector of $\mathcal{C}_h$ associated with the (approximate) eigenvalue $2\ell^s$. Lemma 2.5.2 yields

$$p_{\mathcal{C}_h}(2\ell^s) \equiv 0 \pmod{\ell^{n-\alpha}}.$$

If we let $\pm\mu$ denote the eigenvalues of h, then $p'_{\mathcal{C}_h}(t) = (t^2 - 4\mu^2) + 2t^2$, and evaluating at $2\ell^s$ we find

$$p'_{\mathcal{C}_h}(2\ell^s) = 4(\ell^{2s} - \mu^2) + 8\ell^{2s} = \frac{p_{\mathcal{C}_h}(2\ell^s)}{2\ell^s} + 8\ell^{2s}.$$

To estimate the ℓ-adic valuation of this last expression simply observe that

$$v_\ell\left(\frac{p_{\mathcal{C}_h}(2\ell^s)}{2\ell^s}\right) = v_\ell\left(p_{\mathcal{C}_h}(2\ell^s)\right) - v_\ell(2) - s \geq n - \alpha - v - s > 3v + 2s,$$

so $v_\ell\left(p'_{\mathcal{C}_h}(2\ell^s)\right) = v_\ell\left(8\ell^{2s}\right) = 3v + 2s$. By Hensel's lemma (lemma 2.5.1), $p_{\mathcal{C}_h}(t)$ admits a root $\lambda \in \mathbb{Z}_\ell$ such that

$$v_\ell(\lambda - 2\ell^s) \geq v_\ell(p_{\mathcal{C}_h}(2\ell^s)) - v_\ell(p'_{\mathcal{C}_h}(2\ell^s)) \geq n - \alpha - 2s - 3v > 2s + 1.$$

Note that λ cannot be zero, because clearly $v_\ell(0 - 2\ell^s) = v + s$ is strictly smaller than $v_\ell(\lambda - 2\ell^s)$. It follows that λ is one of the other two roots of $p_{\mathcal{C}_h}(t)$, namely $\pm 2\mu$, and hence

$$\pm\mu = \pm\frac{1}{2}\left(2\ell^s + O(n - \alpha - 2s - 3v)\right) = \pm\ell^s(1 + O(n - \alpha - 3s - 4v)).$$

To sum up, the two eigenvalues of h belong to $\mathbb{Z}_\ell$ and are of the form $\pm\ell^s + O(n - \alpha - 2s - 4v)$ (and in particular of the form $\pm\ell^s + O(s + 4)$). Let μ_+ be the one of the form $\ell^s + O(n - \alpha - 2s - 4v)$ and $v_+ \in \mathbb{Z}_\ell^2$ a corresponding eigenvector, normalized in such a way that at least one of the two coordinates is an ℓ-adic unit. Set furthermore $v_- = yv_+$.
As anticipated, our next objective is to describe the action of x, y, h on $v_\pm$. We expect v_+ to be annihilated by x and v_- to be an eigenvector for h that is annihilated by y: of course this is not

going to be exactly true at all orders, but only up to a certain error term that depends on n, α and s. Let β be the minimal valuation of the coordinates of xv_+: this is a number we want to show to be large.

The idea is that if xv_+ were not very close to zero, then it would be an eigenvector of h associated with an eigenvalue that h does not possess. Note that

$$h(xv_+) \equiv [h,x]v_+ + xhv_+ \equiv (2\ell^s + \mu_+)xv_+ \pmod{\ell^n},$$

so by lemma 2.5.4 either h has an eigenvalue ξ such that $v_\ell(\xi - (\mu_+ + 2\ell^s)) \geq 3 + v_\ell(\mu_+ + 2\ell^s) \geq s+3$ or $\beta \geq n - 2(2 + v_\ell(\mu_+ + 2\ell^s))$. Note now that we cannot be in the first case: indeed h would have an eigenvalue of the form $3\ell^s + O(s+3)$, but we have already seen that the eigenvalues of h are $\pm\ell^s + O(s+4)$, contradiction. Hence we are in the second situation, and furthermore $v_\ell(\mu_+ + 2\ell^s) \leq s+1$: hence $\beta \geq n - 2(2 + v_\ell(\mu_+ + 2\ell^s)) \geq n - 2s - 6$, and by definition of β this means $xv_+ \equiv 0 \pmod{\ell^{n-2(s+3)}}$. Next we compute

$$\begin{aligned} hv_- &= hyv_+ \\ &= [h,y]v_+ + yhv_+ \\ &= -2\ell^s \cdot yv_+ + y(\mu_+ v_+) + O(n) \\ &= (\mu_+ - 2\ell^s)v_- + O(n) \\ &= (-\ell^s + O(n - \alpha - 2s - 4v))v_- \\ &= -\ell^s v_- + O(n - \alpha - 2s - 4v), \end{aligned} \tag{2.2}$$

$$\begin{aligned} xv_- &= xyv_+ \\ &= [x,y]v_+ + yxv_+ \\ &= \ell^s hv_+ + O(n - 2(s+3)) \\ &= \ell^s \mu_+ v_+ + O(n - 2(s+3)) \\ &= \ell^s \left(\ell^s + O(n - \alpha - 2s - 4v)\right) v_+ + O(n - 2(s+3)) \\ &= \ell^{2s} v_+ + O(n - \alpha - 2(s+3)); \end{aligned} \tag{2.3}$$

this settles the question of the action of h and x on v_-. We are left with showing that v_- is (approximately) killed by y:

$$\begin{aligned} h \cdot yv_- &= [h,y]v_- + yhv_- \\ &= -2\ell^s \cdot yv_- + y\left((-\ell^s) + O(n - \alpha - 2s - 4v)\right) v_- + O(n) \\ &= -3\ell^s yv_- + O(n - \alpha - 2s - 4v), \end{aligned}$$

so that yv_- is an (approximate) eigenvector of h, associated with the (approximate) eigenvalue $-3\ell^s$. Let γ be minimal among the valuations of the coefficients of yv_-. Apply lemma 2.5.4: either $\gamma \geq n - \alpha - 2s - 4v - 2(2 + v_\ell(-3\ell^s)) \geq n - \alpha - 4s - 4v - 6$ or h has an eigenvalue ν satisfying $v_\ell(\nu + 3\ell^s) \geq v_\ell(-3\ell^s) + 3 \geq s+3$. This second possibility contradicts what we have already proven on the eigenvalues of h, hence $\gamma \geq n - \alpha - 4s - 4v - 6$, that is to say $yv_- = O(n - \alpha - 4s - 4v - 6)$. Putting it all together, we have proved that up to an error of order $\ell^{n-\alpha-4s-4v-6}$ we have

$$xv_+ = 0,\ yv_+ = v_-,\ hv_+ = \ell^s v_+,\ xv_- = \ell^{2s} v_+,\ yv_- = 0,\ hv_- = -\ell^s v_-.$$

Write $\overline{x}$ (resp. $\overline{y}, \overline{h}$) for $\ell^s \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (resp. $\ell^s \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ell^s \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$) and consider the matrix $\tilde{M}$ whose columns are given by $\ell^s v_+$ and v_-. The above relations may be stated more compactly as

$$\tilde{M}\overline{x} = x\tilde{M},\ \tilde{M}\overline{y} = y\tilde{M},\ \tilde{M}\overline{h} = h\tilde{M} \tag{2.4}$$

modulo $\ell^{n-\alpha-4s-4v-6}$. Let δ be minimal among the valuations of the coefficients of $\tilde{M}$: by construction, at least one of the coordinates of v_+ is an ℓ-adic unit, so $\delta \leq s$. Set $M = \ell^{-\delta}\tilde{M}$. Dividing equations (2.4) by ℓ^δ we see that M satisfies analogous equations up to error terms of order $n-\alpha-5s-4v-6$, and by construction at least one of the coefficients of M is an ℓ-adic unit. Let now g be any element of L_1. The matrix $\ell^s g$ belongs to $\ell^s \mathfrak{sl}_2(\mathbb{Z}_\ell)$, so it is a linear combination of $\overline{x}, \overline{y}, \overline{h}$ with coefficients in $\mathbb{Z}_\ell$. Write $\ell^s g = \lambda_1\overline{x} + \lambda_2\overline{y} + \lambda_3\overline{h}$. We have

$$\begin{aligned} \ell^s M g &= M(\ell^s g) \\ &= M(\lambda_1\overline{x} + \lambda_2\overline{y} + \lambda_3\overline{h}) \\ &= (\lambda_1 x + \lambda_2 y + \lambda_3 h)M + O(n-\alpha-5s-4v-6) \\ &= \varphi(\ell^s g)M + O(n-\alpha-5s-4v-6) \\ &= \ell^s\varphi(g)M + O(n-\alpha-5s-4v-6), \end{aligned}$$

so that dividing by ℓ^s we deduce $Mg = \varphi(g)M + O(n-\alpha-6s-4v-6)$ for every $g \in L_1$, which is the first statement in the proposition.

Let us now turn to the statement concerning the determinant. We can assume that v_+ is normalized so that $v_+ = \begin{pmatrix} 1 \\ c \end{pmatrix}$. We also write $v_- = \begin{pmatrix} b \\ d \end{pmatrix}$. It is clear that $v_\ell(\det M) \leq v_\ell(\det \tilde{M})$, and that $\det \tilde{M} = \ell^s \det \begin{pmatrix} 1 & b \\ c & d \end{pmatrix}$, so let us consider $D := v_\ell\left(\det \begin{pmatrix} 1 & b \\ c & d \end{pmatrix}\right)$. Suppose by contradiction $D > 3s+v$; by definition of the determinant we have $d = bc + O(D)$, which implies

$$v_- = \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} b \\ bc + O(D) \end{pmatrix} = bv_+ + O(D).$$

Applying h to both sides of this equality and using equation (2.2) we get

$$\mu_- v_- + O(n-\alpha-2s-4v) = hv_- = h(bv_+ + O(D)) = b\mu_+ v_+ + O(D).$$

Comparing the first coordinate of these vectors we deduce

$$b\mu_- = b\mu_+ + O(\min\{D, n-\alpha-2s-4v\}),$$

hence

$$\mu_- = \mu_+ + O(\min\{D - v_\ell(b), n-\alpha-2s-4v-v_\ell(b)\}). \tag{2.5}$$

Note now that since $d = bc + O(D)$ we have $v_\ell(d) \geq \min\{v_\ell(b), D\}$. Moreover, we see by equation (2.3) that $xv_- = \ell^{2s}v_+ + O(n-\alpha-2(s+3))$, and since the right hand side does not vanish modulo ℓ^{2s+1} (since $n-\alpha-2(s+3) > 2s+1$ and $\ell^{2s}v_+ = \begin{pmatrix} \ell^{2s} \\ \ell^{2s}c \end{pmatrix}$) we deduce that $\min\{v_\ell(b), v_\ell(d)\} \leq 2s$. Let us show that we also have $v_\ell(b) \leq 2s$. Suppose that $v_\ell(b) \geq 2s+1$: then

$$v_\ell(d) \geq \min\{v_\ell(b), D\} \geq \min\{2s+1, 3s+v+1\} \geq 2s+1,$$

which implies $\min\{v_\ell(b), v_\ell(d)\} \geq 2s+1$ and contradicts what we just proved. Therefore $v_\ell(b) \leq 2s$, hence equation (2.5) implies $\mu_- = \mu_+ + O(D-2s)$: notice that if the minimum in (2.5) were attained for $n-\alpha-2s-4v-v_\ell(b) > 3s+1$ we would have $\ell^s = -\ell^s + O(3s+2)$, a clear contradiction. On the other hand, we know that $\mu_\pm = \pm\ell^s + O(s+4)$, so the above equation implies $2\ell^s + O(s+4) = O(D-2s)$. Hence we have proved $v_\ell(2\ell^s) \geq D-2s$, i.e. $D \leq 3s+v$, a contradiction. It follows, as claimed, that $v_\ell(\det M) \leq v_\ell(\det \tilde{M}) = s+D \leq 4s+v$.
Next we prove the statement concerning traces. Let g be any element of L_1. Setting, for the sake of simplicity, $N = n-\alpha-6s-4v-6$, we have $Mg = \varphi(g)M + O(N)$, so (multiplying on the left by the adjugate M^* of M) we deduce $\det(M)g = M^*\varphi(g)M + O(N)$. Didiving through by $\det(M)$ we have $g = M^{-1}\varphi(g)M + O(N-(4s+v))$; note that this equality would a priori only hold in $\mathfrak{sl}_2(\mathbb{Q}_\ell)$, but since both g and the error term are ℓ-integral we necessarily also have $M^{-1}\varphi(g)M \in \mathfrak{sl}_2(\mathbb{Z}_\ell)$. Squaring and taking traces then yields $\operatorname{tr}\left(g^2\right) = \operatorname{tr}\left[\left(M^{-1}\varphi(g)M\right)^2\right] + O(N-(4s+v))$, i.e.

$$\operatorname{tr}\left(g^2\right) = \operatorname{tr}\left(\varphi(g)^2\right) + O(N-(4s+v))$$

as claimed. Finally, essentially the same argument shows the last two statements: we can multiply the congruence $Mg_1 \equiv \varphi(g_1)M \pmod{\ell^N}$ on the right (resp. left) by M^* and divide by $\det M$ to get

$$Mg_1M^{-1} \equiv \varphi(g_1) \pmod{\ell^{N-4s-v}}, \quad g_1 \equiv M^{-1}\varphi(g_1)M \pmod{\ell^{N-4s-v}}.$$

□

2.6 Products of two curves

Let E_1, E_2 be two elliptic curves over K and ℓ be a prime number. To study the Galois representation attached to $E_1 \times E_2$ we are going to pass to a suitable extension of K over which the study of the Lie algebra of G_ℓ (the image of $\operatorname{Gal}(\overline{K}/K)$ inside $\operatorname{Aut} T_\ell(E_1) \times \operatorname{Aut} T_\ell(E_2) \cong \mathrm{GL}_2(\mathbb{Z}_\ell)^2$) is sufficient to yield information on G_ℓ itself. Before doing this, however, we need to dispense with some necessary preliminaries. Let $G_{\ell,1}, G_{\ell,2}$ be the two projections of G_ℓ onto the two factors $\mathrm{GL}_2(\mathbb{Z}_\ell)$, and m_1, m_2 be integers such that $\mathcal{B}_\ell(m_i)$ is contained in $G_{\ell,i}$ for $i = 1, 2$.
Suppose for the moment that ℓ is odd. We want to apply theorem 2.4.2, so for the whole section (up until the very last proposition) we make the following

Assumption. If ℓ is odd, G_ℓ does not contain $\mathcal{B}_\ell\left(20\max\{m_1,m_2\}, 20\max\{m_1,m_2\}\right)$.

Under this assumption, we define K_ℓ to be the extension of K associated with the following closed subgroups of G_ℓ:

$$\begin{cases} \ker\left(G_2 \to \mathrm{GL}_2(\mathbb{Z}/8\mathbb{Z})^2\right), & \text{if } \ell = 2 \\ H_\ell, & \text{if } \ell \neq 2, \end{cases}$$

where H_ℓ is the group given by an application of theorem 2.4.2 under our assumption. Note that the degree $[K_2 : K]$ is at most $3^2 2^{16}$, that is to say the order of

$$\left\{(x,y) \in \mathrm{GL}_2(\mathbb{Z}/8\mathbb{Z})^2 \;\middle|\; \det x = \det y\right\},$$

whereas $[K_\ell : K]$ is uniformly bounded by $2 \cdot 48^2$ for $\ell \neq 2$. Note that H_ℓ is by construction the image of $\operatorname{Gal}(\overline{K_\ell}/K_\ell)$ in $\operatorname{Aut} T_\ell(E_1) \times \operatorname{Aut} T_\ell(E_2) \cong \mathrm{GL}_2(\mathbb{Z}_\ell)^2$; we write $H_{\ell,1}, H_{\ell,2}$ for its two

projections on the two factors $\mathrm{GL}_2(\mathbb{Z}_\ell)$. Furthermore, we let n_1, n_2 be integers such that $H_{\ell,1}, H_{\ell,2}$ respectively contain $\mathcal{B}_\ell(n_1), \mathcal{B}_\ell(n_2)$. Notice that if $\ell = 2$ we have $n_1, n_2 \geq 2$; on the other hand, for $\ell = 3$ or t we explicitly demand that $n_1, n_2 \geq 1$ (the groups H_3 and H_5 as constructed in chapter 8 will automatically satisfy this inequality).

Remark 2.6.1. Note that if $m_1, m_2 > 0$ we can in fact take $n_1 = m_1, n_2 = m_2$ unless $\ell \leq 3$: indeed for primes $\ell \geq 5$ the index of H_ℓ in G_ℓ is not divisible by ℓ, so for any positive value of n the (pro-ℓ) group $\mathcal{B}_\ell(n)$ is contained in $H_{\ell,i}$ if and only if it is contained in $G_{\ell,i}$.

Let $L \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)^{\oplus 2}$ (resp. $L_1, L_2 \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)$) be the Lie algebra of H_ℓ (resp. $H_{\ell,1}, H_{\ell,2}$). Choose a basis of L of the form $(a_1, b_1), (a_2, b_2), (a_3, b_3), (0, y_1), (0, y_2), (0, y_3)$. Such a basis clearly exists. Since by our assumption $H_{\ell,1} \supseteq \mathcal{B}_\ell(n_1)$ we have $L_1 \supseteq \ell^{n_1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.
Also note that $(0, y_1), (0, y_2), (0, y_3)$ span a Lie-subalgebra: indeed $[(0, y_i), (0, y_j)] = (0, [y_i, y_j])$ must be a linear combination with $\mathbb{Z}_\ell$-coefficients of the basis elements; however, since a_1, a_2, a_3 are linearly independent over $\mathbb{Z}_\ell$, we deduce that this commutator is a linear combination of $(0, y_1), (0, y_2), (0, y_3)$, so that these three elements do indeed span a Lie algebra, which we call L_3. Note that L_3 can equivalently be described as the kernel of the projection from $L \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$ to the first copy of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. We shall interchangeably think of L_3 as being a subaglebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ or of $\mathfrak{sl}_2(\mathbb{Z}_\ell)^{\oplus 2}$, by identifying it with its projection on the second factor $\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

Lemma 2.6.2. *$L_3 \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)$ is stable under conjugation by $\mathcal{B}_\ell(n_2)$.*

Proof. For the proof we consider L_3 as a subalgebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)^{\oplus 2}$.
Take any element $l \in L_3$: it is the limit of a certain sequence $l_n = \sum_{i=1}^{n} \lambda_{n,i}\Theta(g_{n,i})$ for certain $g_{n,i} \in H_\ell$. For any $g \in B_\ell(n_2)$ there exists a certain $h \in H_{\ell,1}$ such that (h, g) is in H_ℓ. We have

$$\begin{aligned}(h,g)^{-1}l_n(h,g) &= \sum_{i=1}^{n} \lambda_{n,i}(h,g)^{-1}\Theta(g_{n,i})(h,g) = \sum_{i=1}^{n} \lambda_{n,i}(h,g)^{-1}\left(g_{n,i} - \frac{\mathrm{tr}(g_{n,i})}{2}\,\mathrm{Id}\right)(h,g)\\ &= \sum_{i=1}^{n} \lambda_{n,i}\left((h,g)^{-1}g_{n,i}(h,g) - \frac{\mathrm{tr}((h,g)^{-1}g_{n,i}(h,g))}{2}\,\mathrm{Id}\right)\\ &= \sum_{i=1}^{n} \lambda_{n,i}\Theta((h,g)^{-1}g_{n,i}(h,g)) \in \langle\Theta(H_\ell)\rangle,\end{aligned}$$

so the sequence $\left((h,g)^{-1}l_n(h,g)\right)_{n\geq 0}$ is in L, and by continuity of conjugation tends to the element $(h,g)^{-1}l(h,g)$ of L. Now if we write $l = (l^{(1)}, l^{(2)}) = (0, l^{(2)})$ we have

$$(h,g)^{-1}l(h,g) = (h,g)^{-1}(0,l^{(2)})(h,g) = (0, g^{-1}l^{(2)}g) \in L,$$

and since L_3 is exactly the sub-algebra given by the elements whose first coordinate vanishes the claim is proved. □

Lemma 2.6.3. *Fix an integer t, and suppose that at least one among y_1, y_2, y_3 is not zero modulo ℓ^{t+1}: then L_3 contains $\ell^{t+4n_2+4v}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.*

Proof. Apply lemma 2.4.4 with $s = n_2$ (recalling that $n_2 \geq 1$ if $\ell = 3$ or 5). □

Our task is therefore to bound the values of t for which the y_i's all vanish modulo ℓ^t. If at least one among y_1, y_2, y_3 does not vanish modulo ℓ^{n_2+1} we are done, so we can assume without loss of

generality that $y_i \equiv 0 \pmod{\ell^t}$ for a certain $t \geq n_2 + 1$. If this is the case, then none of b_1, b_2, b_3 can be zero modulo ℓ^t, for otherwise L_2 could not contain $\ell^{n_2}\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Even more, the b_i's must generate $\ell^{n_2}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

Denote by $\varphi : L_1 \to L_2$ the only $\mathbb{Z}_\ell$-linear map sending a_i to b_i for $i = 1, 2, 3$. For two indices j, k write $[a_j, a_k] = \sum_{i=1}^3 \mu_i^{(j,k)} a_i$. There exist scalars $\nu_i^{(j,k)}$ such that

$$[(a_j, b_j), (a_k, b_k)] = \sum_{i=1}^3 \mu_i^{(j,k)}(a_i, b_i) + \sum_{i=1}^3 \nu_i^{(j,k)}(0, y_i),$$

and reducing the second coordinate of this equation modulo ℓ^t gives

$$\begin{aligned}
[\varphi(a_j), \varphi(a_k)] = [b_j, b_k] &\equiv \sum_{i=1}^3 \mu_i^{(j,k)} b_i \\
&\equiv \sum_{i=1}^3 \mu_i^{(j,k)} \varphi(a_i) \\
&\equiv \varphi\left(\sum_{i=1}^3 \mu_i^{(j,k)} a_i\right) \\
&\equiv \varphi\left([a_j, a_k]\right) \pmod{\ell^t}.
\end{aligned}$$

We want to apply proposition 2.5.5 to φ. We claim that, in the notation of that proposition, we can take $\alpha \leq n_2 + n_1$. Since b_1, b_2, b_3 generate $\ell^{n_2}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, a linear combination $\lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 b_3$ vanishes modulo $\ell^{n_1+n_2+1}$ only if $\lambda_1, \lambda_2, \lambda_3$ all vanish modulo ℓ^{n_1+1}. Now since a_1, a_2, a_3 generate $\ell^{n_1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$ we can choose scalars $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{Z}_\ell$ such that $\ell^{n_1}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3$, and clearly at least one among λ_1, λ_2, λ_3 is nonzero modulo ℓ^{n_1+1}. It follows that

$$\varphi\left(\ell^{n_1}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = \varphi\left(\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3\right) = \sum_{i=1}^3 \lambda_i b_i$$

is nonzero modulo $\ell^{n_1+n_2+1}$ as claimed, and a perfectly analogous argument applies to the image of $\ell^{n_1}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Also note that by construction of φ and by our assumption on t we have

$$(l_1, l_2) \in L(H_\ell) \Rightarrow l_2 \equiv \varphi(l_1) \pmod{\ell^t}.$$

Set $T = t - 11n_1 - n_2 - 5v - 6$. By proposition 2.5.5, there is a matrix $M \in M_2(\mathbb{Z}_\ell)$ such that:

1. $\operatorname{tr} l_1^2 \equiv \operatorname{tr}(\varphi(l_1)^2) \equiv \operatorname{tr} l_2^2 \pmod{\ell^T} \quad \forall (l_1, l_2) \in L(H_\ell)$;
2. for all $(l_1, l_2) \in L(H_\ell)$ we have $l_2 \equiv M \cdot l_1 \cdot M^{-1} \pmod{\ell^T}$ and $M^{-1} \cdot l_2 \cdot M \equiv l_1 \pmod{\ell^T}$;
3. at least one of the coefficients of M is an ℓ-adic unit.

Take any element $(g_1, g_2) \in H_\ell$. By our choice of K_ℓ, we know that the determinant of g_1 is a square in $\mathbb{Z}_\ell$, so we can choose a square root of $\det g_1$ and write

$$(g_1, g_2) = \sqrt{\det g_1}(g_1', g_2')$$

for a certain $(g_1', g_2') \in \mathrm{SL}_2(\mathbb{Z}_\ell)$. The image (l_1, l_2) of (g_1', g_2') via Θ_2 differs from $\Theta_2(g_1, g_2)$ by a scalar multiple, so it lies again in $L(H_\ell)$. By definition, there exists a pair $(\lambda_1, \lambda_2) \in \mathbb{Z}_\ell^2$ such that

$$(g_1', g_2') = (\lambda_1, \lambda_2) \cdot \mathrm{Id} + (l_1, l_2), \tag{2.6}$$

and we wish to show that λ_1 is congruent to λ_2 modulo a large power of ℓ. We begin by discussing the case of odd ℓ. Squaring equation (2.6) we obtain

$$\left(\left(g_1'\right)^2, \left(g_2'\right)^2\right) = (\lambda_1^2 \cdot \mathrm{Id} + l_1^2 + 2\lambda_1 l_1, \lambda_2^2 \cdot \mathrm{Id} + l_2^2 + 2\lambda_2 l_2).$$

Now the left hand side is simply $\dfrac{1}{\det g_1}\left(g_1^2, g_2^2\right)$, an element of H_ℓ up to scalar multiples. The image of this matrix through Θ_2 is then an element of $L(H_\ell)$, so applying Θ_2 to the right hand side of the previous equation we get

$$\left(\Theta_1(l_1^2) + 2\lambda_1 l_1, \Theta_1(l_2^2) + 2\lambda_2 l_2\right) \in L(H_\ell), \tag{2.7}$$

which implies

$$\Theta_1(l_2^2) + 2\lambda_2 l_2 \equiv M\left(\Theta_1(l_1^2) + 2\lambda_1 l_1\right) M^{-1} \pmod{\ell^T}$$

and, using properties (1) and (2) above,

$$2\lambda_1 l_2 \equiv M\left(2\lambda_1 l_1\right) M^{-1} \equiv 2\lambda_2 l_2 \pmod{\ell^T}.$$

If l_2 has at least one coordinate not divisible by ℓ, this last equation implies $\lambda_1 \equiv \lambda_2 \pmod{\ell^T}$. If not, then g_2' reduces modulo ℓ to a multiple of the identity (cf. equation (2.6)). Moreover, as $\det(g_2') = 1$, we have in particular

$$1 = \det(\lambda_2 \,\mathrm{Id} + l_2) = \lambda_2^2 - \frac{\mathrm{tr}\left(l_2^2\right)}{2},$$

from which we find

$$\lambda_2 = \pm\sqrt{1 + \frac{\mathrm{tr}(l_2^2)}{2}},$$

where the series converges since l_2 is trivial modulo ℓ. Symmetrically we prove that either the congruence $\lambda_1 \equiv \lambda_2 \pmod{\ell^T}$ holds or else l_1 is trivial modulo ℓ and

$$\lambda_1 = \pm\sqrt{1 + \frac{\mathrm{tr}(l_1^2)}{2}}.$$

Suppose then l_1, l_2 to be both trivial modulo ℓ. As $\mathrm{tr}\left(l_1^2\right) \equiv \mathrm{tr}\left(l_2^2\right) \pmod{\ell^T}$, it follows that λ_1 and λ_2 are congruent modulo ℓ^T as long as λ_1 and λ_2 have the same reduction modulo ℓ. But g_1', g_2' reduce to diagonal matrices $\mathrm{diag}\,(\lambda_i, \lambda_i)$ in $\mathrm{SL}_2(\mathbb{F}_\ell)$, so $\lambda_1 \equiv \lambda_2 \pmod{\ell^T}$ if and only if g_1', g_2' have the same reduction modulo ℓ, and this is exactly one of the properties of H_ℓ given by theorem 2.4.2. If, on the other hand, $\ell = 2$, then l_1, l_2 vanish modulo 4 by construction and the same argument as above shows that

$$\lambda_i = \pm\sqrt{1 + \frac{\mathrm{tr}(l_i^2)}{2}},\ i = 1, 2. \tag{2.8}$$

Given that $2\lambda_i \equiv \mathrm{tr}\,(g_i') \equiv 2 \pmod 8$ by our construction of H_ℓ, it follows that $\lambda_1 \equiv \lambda_2 \equiv 1 \pmod 4$, so the sign in equation (2.8) must be a plus and $\lambda_1 \equiv \lambda_2 \pmod{2^{T-2}}$. Using this information in equation (2.6) we have thus proved

Lemma 2.6.4. *There exists a matrix $M \in M_2(\mathbb{Z}_\ell)$ such that, for every element $(g_1, g_2) \in H_\ell$, the congruence $g_2 \equiv M g_1 M^{-1} \pmod{\ell^{T-2v}}$ holds.*

Set now $H := T - 2v$ and choose any $w \in E_1[\ell^H]$: as $\ell^H w = 0$, for every $(g_1, g_2) \in H_\ell$ we have

$$Mg_1 w = Mg_1 M^{-1} Mw = (g_2 M + O(\ell^H))w = g_2 Mw,$$

so the subgroup

$$\Gamma = \left\{(w, Mw) \in E_1[\ell^H] \times E_2[\ell^H] \;\middle|\; w \in E_1[\ell^H]\right\}$$

is defined over K_ℓ: indeed for any $(g_1, g_2) \in H_\ell$ we have

$$(g_1, g_2) \cdot (w, Mw) = (g_1 w, g_2 M w) = (g_1 w, M g_1 w).$$

Thus the abelian variety $A^* = E_1 \times E_2/\Gamma$ is defined over K_ℓ, and we have an isogeny $E_1 \times E_2 \to A^*$ of degree $|E_1[\ell^H]| = \ell^{2H}$; on the other hand, we also have an isogeny $A^* \to E_1 \times E_2$ of degree b dividing $b_0(E_1 \times E_2/K_\ell)$, and the composition of the two is an endomorphism of $E_1 \times E_2$ that kills Γ. Here we use the crucial fact that at least one of the coefficients of M is an ℓ-adic unit to deduce that the projection of Γ on E_2 contains points of exact order ℓ^H, so the endomorphism of $E_1 \times E_2$ killing Γ must be of the form $\begin{pmatrix} \ell^H e_1 & 0 \\ 0 & \ell^H e_2 \end{pmatrix}$, of degree $e_1^2 e_2^2 \ell^{4H}$. It follows that $e_1^2 e_2^2 \ell^{4H} = \ell^{2H} b$, hence $2H \leq v_\ell(b_0(E_1 \times E_2/K_\ell))$ and $2t \leq v_\ell(b_0(E_1 \times E_2/K_\ell)) + 2(11n_1 + n_2 + 7v + 6)$. This inequality is certainly not satisfied if we take $t = \left\lfloor \frac{v_\ell(b_0(E_1 \times E_2/K_\ell))}{2} \right\rfloor + 11n_1 + n_2 + 7v + 7$, so for this value of t the Lie algebra L_3 does not vanish modulo ℓ^t. Lemma 2.6.3 then shows that L_3 contains $\ell^{f_1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, where $f_1 = \left\lfloor \frac{v_\ell(b_0(E_1 \times E_2/K_\ell))}{2} \right\rfloor + 11n_1 + 5n_2 + 11v + 7$, and therefore $L(H_\ell)$ contains $0 \oplus \ell^{f_1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Swapping the roles of E_1 and E_2 we deduce that $L(H)$ contains $\ell^f \mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^f \mathfrak{sl}_2(\mathbb{Z}_\ell)$, where now

$$f = \left\lfloor \frac{v_\ell(b_0(E_1 \times E_2/K_\ell))}{2} \right\rfloor + 16 \max\{n_1, n_2\} + 11v + 7.$$

Proposition 2.6.5. *Let E_1, E_2 be elliptic curves over K that are not isogenous over $\overline{K}$ and do not admit complex multiplication over $\overline{K}$. Let ℓ be a prime number.*
Suppose the image of $\mathrm{Gal}(\overline{K_\ell}/K_\ell) \to \mathrm{Aut}(T_\ell(E_i))$ contains $\mathcal{B}_\ell(n_i)$ for $i = 1, 2$ (where $n_i \geq 2$ for $\ell = 2$ and $n_i \geq 1$ for $\ell = 3$ or 5). Let f be given by the formula above. If ℓ is odd, the image G_ℓ of $\mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(T_\ell(E_1)) \times \mathrm{Aut}(T_\ell(E_2))$ contains $\mathcal{B}_\ell(4f+4) \times \mathcal{B}_\ell(4f+4)$; if $\ell = 2$, the image G_2 of $\mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(T_2(E_1)) \times \mathrm{Aut}(T_2(E_2))$ contains

$$\mathcal{B}_2(12(f + 17\max\{n_1, n_2\} + 13) + 1, 12(f + 17\max\{n_1, n_2\} + 13) + 1).$$

Proof. For $\ell = 2$ the result follows at once from theorem 2.4.3. For odd ℓ, and under the assumption we made at the beginning of this section, the result similarly follows from property $(*)$ of H_ℓ given in theorem 2.4.2 and the fact that clearly $2f + 4 > 8\max\{n_1, n_2\}$. On the other hand, if our assumption is *false*, then G_ℓ contains $\mathcal{B}_\ell(20\max\{n_1, n_2\}, 20\max\{n_1, n_2\})$ (note that we can assume $m_1 \leq n_1, m_2 \leq n_2$ without loss of generality), which is stronger than what we are claiming. □

2.7 Conclusion

Consider again the case of two elliptic curves E_1, E_2 defined over K, non-isogenous over $\overline{K}$ and such that $\mathrm{End}_{\overline{K}}(E_i) = \mathbb{Z}$. Let $\mathcal{P}$ be the set of primes ℓ for which G_ℓ does not contain $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$. Rewriting Proposition 1 of [71] in terms of the function b_0 of definition 2.2.4 we get:

Lemma 2.7.1. *Let ℓ be a prime. If ℓ does not divide the product*

$$30 b_0(E_1/K; 60) b_0\left(E_1^2/K; 2\right) b_0(E_2/K; 60) b_0\left(E_2^2/K; 2\right) b_0(E_1 \times E_2/K; 2),$$

then ℓ is not in $\mathcal{P}$.

Proof. Lemma 1.8.2 implies that for a prime ℓ that does not divide

$$b_0(E_1/K;60)b_0\left(E_1^2/K;2\right)b_0(E_2/K;60)b_0\left(E_2^2/K;2\right)$$

both projections of $G_\ell(\ell)$ on the two factors $\mathrm{GL}_2(\mathbb{F}_\ell)$ contain $\mathrm{SL}_2(\mathbb{F}_\ell)$. Under this hypothesis, the proof of [71, Proposition 1] shows that $G_\ell(\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)^2$ unless $\ell^2 \mid b_0(E_1 \times E_2/K;2)$. Finally, a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)^2$ whose projection modulo ℓ contains $\mathrm{SL}_2(\mathbb{F}_\ell)^2$ contains all of $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$, if $\ell \geq 5$ (this is well-known; see for example [113, Proposition 4.2]). □

Corollary 2.7.2. *The inequality*

$$\prod_{\ell\in\mathcal{P}} \ell \leq 30b_0(E_1/K;60)b_0\left(E_1^2/K;2\right)b_0(E_2/K;60)b_0\left(E_2^2/K;2\right)b_0(E_1\times E_2/K;2)$$

holds.

Let now ℓ be a prime different from 2 and 3. For $j = 1, 2$ set

$$D_j(\infty) = b_0(E_j/K;120)^5 b_0(E_j^2/K;2).$$

As ℓ is odd, by corollary 1.7.6 we see that $G_{\ell,j}$ contains

$$\mathcal{B}_\ell\left(16v_\ell(D_j(\infty)) + 12\right),$$

hence the same is true for $H_{\ell,j}$, cf. remark 2.6.1. Therefore – in the notation of the previous section – we can take $n_j = n_j(\ell) = 16v_\ell(D_j(\infty)) + 12$. On the other hand, for $\ell = 3$ we apply theorem 1.7.5 directly to E_j/K_3 (notice that our present K_3 satisfies the same hypotheses as the field noted K_3 in chapter 1) and see that we can take

$$n_j(3) = 16v_3\left(b_0(E/K_3)^5 b_0(E^2/K_3)\right) + 12 \leq 16v_3(D_j(\infty)) + 12;$$

similarly, for $\ell = 2$ we can take $n_j(2) = 48v_2\left(b_0(E_j/K_2)^5 b_0(E_j^2/K_2)\right) + 38$.
Applying proposition 2.6.5 with these values of n_j we get:

Lemma 2.7.3. *Let ℓ be a prime. The group G_ℓ contains $\mathcal{B}_\ell(f(\ell), f(\ell))$, where $f(\ell)$ is given by*

$$f(\ell) = 2v_\ell(b_0(E_1 \times E_2/K; 2\cdot 48^2)) + 2^{10}\max\left\{v_\ell(D_1(\infty)), v_\ell(D_2(\infty))\right\} + 10^3$$

for odd ℓ and

$$f(2) = 6v_2(b_0(E_1 \times E_2/K_2)) + 10^4 \max_j\left\{v_2\left(b_0(E_j/K_2)^5 b_0(E_j^2/K_2)\right)\right\} + 10^4$$

for $\ell = 2$.

Using the very same argument as in chapter 1 (paragraph 1.8), and some very crude estimates, we deduce

Proposition 2.7.4. *G_∞ contains a subgroup S of the form $S = \prod_\ell S_\ell$, where each S_ℓ coincides with $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$ except for the finitely many primes that are in $\mathcal{P}$, for which $S_\ell = \mathcal{B}_\ell(f(\ell), f(\ell))$. The index of S in $\mathrm{SL}_2(\hat{\mathbb{Z}})$ is bounded by $b(E_1 \times E_2/K; 2\cdot 48^2)^{12000}$.*

We finally come to the adelic estimate for an arbitrary number of curves:

Theorem 2.7.5. *Let $E_1, \ldots, E_n$, $n \geq 2$, be elliptic curves defined over K, pairwise non-isogenous over $\overline{K}$. Suppose that $\mathrm{End}_{\overline{K}}(E_i) = \mathbb{Z}$ for $i = 1, \ldots, n$. Then G_∞ has index at most*

$$(8\zeta(2))^{n(n-1)} \cdot [K:\mathbb{Q}] \cdot \max_{i\neq j} b\left(E_i \times E_j/K; 2\cdot 48^2\right)^{6000n(n-1)}$$

in

$$\left\{(x_1,\dots,x_n)\in \mathrm{GL}_2(\hat{\mathbb{Z}})^n \;\middle|\; \det x_i = \det x_j \quad \forall i,j\right\}.$$

Proof. The exact sequence

$$1\to G_\infty\cap \mathrm{SL}_2(\hat{\mathbb{Z}})^n \to G_\infty \xrightarrow{\det} \hat{\mathbb{Z}}^\times \to \frac{\hat{\mathbb{Z}}^\times}{\det\circ\rho_\infty \mathrm{Gal}(\overline{K}/K)}\to 1$$

and the fact that $\left|\dfrac{\hat{\mathbb{Z}}^\times}{\det\circ\rho_\infty \mathrm{Gal}(\overline{K}/K)}\right| \le [K:\mathbb{Q}]$ (cf. proposition 1.8.1) show that it is enough to prove that the index of $G_\infty\cap \mathrm{SL}_2(\hat{\mathbb{Z}})^n$ inside $\mathrm{SL}_2(\hat{\mathbb{Z}})^n$ is bounded by

$$(8\zeta(2))^{n(n-1)}\cdot \max_{i\neq j} b(E_i\times E_j/K; 2\cdot 48^2)^{6000n(n-1)}.$$

Set $G = G_\infty\cap \mathrm{SL}_2(\hat{\mathbb{Z}})$. For every pair E_i, E_j of curves, we get from proposition 2.7.4 a subgroup $S^{(i,j)}$ of

$$\mathrm{SL}_2(\hat{\mathbb{Z}})^2 \subseteq \prod_\ell \mathrm{Aut}\,(T_\ell(E_i))\times \prod_\ell \mathrm{Aut}\,(T_\ell(E_j))$$

that satisfies all the requirements of corollary 2.3.3, and the theorem follows from this corollary upon noticing that the index of $S^{(i,j)}$ in $\mathrm{SL}_2(\hat{\mathbb{Z}})^2$ is bounded by $b(E_i\times E_j/K; 2\cdot 48^2)^{12000}$. □

Chapter 3

Abelian surfaces & GL_2-varieties

3.1 Introduction

The purpose of this work is the study of Galois representations attached to abelian surfaces over number fields. Throughout the chapter, the letters K and A will respectively denote a number field and a 2-dimensional abelian variety ('surface') defined over K, and the letter ℓ will be reserved for prime numbers. The representations we examine are those given by the natural action of $\mathrm{Gal}(\overline{K}/K)$ on the various Tate modules of A (denoted by $T_\ell(A)$), and the problem we study is that of describing the image G_{ℓ^∞} of $\mathrm{Gal}(\overline{K}/K)$ in $\mathrm{Aut}\,(T_\ell(A))$.
In a sense that will be made precise shortly, we aim to show that this image is as large as it is permitted by some 'obvious' constraints, as soon as ℓ exceeds a certain bound $\ell_0(A, K)$ that we explicitly compute in terms of arithmetical invariants of K and of the semistable Faltings height of A (denoted by $h(A)$). Note that this fact, in its qualitative form, has been known since the work of Serre [118] and Ribet [109]: the novelty of the result we present here lies in its being completely *explicit*. Indeed, to the best of the author's knowledge, before the present work the only paper dealing with the problem of explicit surjectivity results for Abelian surfaces was [48], that only covered the case $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. Unfortunately, the argument of [48] seems to contain a gap, for in his case analysis the author does not include the subgroup of $\mathrm{GSp}_4(\mathbb{F}_\ell)$ arising from the unique 4-dimensional symplectic representation of SL_2 (case 7 in theorem 3.3.2). This is essentially the hardest case, and dealing with it requires nontrivial results of Raynaud on the structure of the action of inertia.
Before stating our main result let us elaborate a little on the 'obvious' conditions that are imposed on G_{ℓ^∞}. On the one hand, the compatibility of the Galois action with the Weil pairing $\langle \cdot, \cdot \rangle$ forces G_{ℓ^∞} to be contained in the group of similitudes with respect to the bilinear form $\langle \cdot, \cdot \rangle$; on the other hand, the action of $\mathrm{Gal}(\overline{K}/K)$ is also compatible with the action of $\mathrm{End}_K(A)$, so that we also know that G_{ℓ^∞} is contained in the centralizer of $\mathrm{End}_K(A)$ inside $\mathrm{Aut}(T_\ell(A))$.

This second condition leads naturally to classifying abelian surfaces according to the structure of $\mathrm{End}_{\overline{K}}(A)$. A study of those rings that appear as endomorphism rings of abelian surfaces (a particular case of the so-called Albert classification, cf. for example [86, p. 203]) leads to the conclusion that only five cases can arise:

1. Type I, trivial endomorphisms: A is absolutely simple and $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$;
2. Type I, real multiplication: A is absolutely simple and $\mathrm{End}_{\overline{K}}(A)$ is an order in a real quadratic field;
3. Type II, quaternionic multiplication: A is absolutely simple and $\mathrm{End}_{\overline{K}}(A)$ is an order in a quaternion division algebra over $\mathbb{Q}$;
4. Type IV, complex multiplication: A is absolutely simple and admits complex multiplication by a quartic CM field;
5. Non-simple case: $A_{\overline{K}}$ is isogenous to the product of two elliptic curves.

We focus here on the first three possibilities. The case of complex multiplication (in arbitrary dimension) is treated in chapter 5, and that of a product of an arbitrary number of elliptic curves without complex multiplication is studied in chapter 2; combining these results, it should also be possible to treat the case of a product of two elliptic curves $E_1 \times E_2$, where E_1 admits complex multiplication and E_2 does not.

3.1.1 Notation and statement of the result

We are interested in the Galois representations attached to A: the natural action of $\mathrm{Gal}(\overline{K}/K)$ on the Tate modules $T_\ell(A)$ gives rise to a family of representations

$$\rho_{\ell^\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(T_\ell(A))$$

which will be our main object of study. We will also need to consider the residual mod-ℓ representations, which we similarly denote by $\rho_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(A[\ell])$, and we write G_{ℓ^∞} (resp. G_ℓ) for the image of ρ_{ℓ^∞} (resp. ρ_ℓ). Most of our estimates will be given in terms of the following function:

Definition 3.1.1. Let K be a number field and A be an abelian variety of dimension g defined over K. Let $\alpha(g) = 2^{10}g^3$ and define

$$b(A/K) = b([K:\mathbb{Q}], g, h(A)) = \left((14g)^{64g^2}[K:\mathbb{Q}]\max\left(h(A), \log[K:\mathbb{Q}], 1\right)^2\right)^{\alpha(g)}.$$

We are now ready to state our main results. Let A/K be an abelian surface, R be its endomorphism ring $\mathrm{End}_{\overline{K}}(A)$, and ℓ be a rational prime.

Theorem 3.1.2. *Suppose that $R = \mathbb{Z}$. The equality $G_{\ell^\infty} = \mathrm{GSp}_4(\mathbb{Z}_\ell)$ holds for every prime ℓ that is not divisible by any place of bad reduction of A, does not ramify in K, and is strictly larger than $b(4 \cdot 1920[K:\mathbb{Q}], 4, 2h(A))^{1/4}$.*

Remark 3.1.3. As it is remarked in chapter 4 (cf. especially remark 4.5.3), the good reduction assumption can be weakened to the assumption that A has *semistable* reduction *at least at one* place of K of characteristic ℓ. Furthermore, by the techniques of that chapter, the bound can be improved to $b(4[K:\mathbb{Q}], 4, 2h(A))^{1/4}$, cf. remark 3.3.20.

For the case of real multiplication, we treat the more general situation of abelian varieties of GL_2-type, namely those abelian varieties A/K such that $\mathrm{End}_{\overline{K}}(A) \otimes \mathbb{Q}$ is a totally real number field whose degree over $\mathbb{Q}$ equals $\dim A$ (such varieties were considered by Ribet in [109]):

Theorem 3.1.4. *Let A/K be an abelian variety of dimension g. Suppose that $R = \mathrm{End}_{\overline{K}}(A)$ is an order in a totally real field E of degree g over $\mathbb{Q}$ and that all endomorphisms of A are defined over K. Let ℓ be a prime unramified both in K and in E and strictly larger than both $b(A/K)^g$ and $b(2[K:\mathbb{Q}], 2\dim(A), 2h(A))^{1/2}$: then we have*

$$G_{\ell^\infty} = \left\{ x \in \mathrm{GL}_2\left(\mathcal{O}_E \otimes \mathbb{Z}_\ell\right) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}.$$

The case of abelian surfaces with real multiplication then follows as an immediate consequence:

Corollary 3.1.5. *Suppose that R is an order in a real quadratic field E and that all endomorphisms of A are already defined over K. Let ℓ be a rational prime, unramified both in K and in E and strictly larger than $b(2[K:\mathbb{Q}], 4, 2h(A))^{1/2}$: then we have*

$$G_{\ell^\infty} = \left\{ x \in \mathrm{GL}_2\left(\mathcal{O}_E \otimes \mathbb{Z}_\ell\right) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}.$$

Remark 3.1.6. When A is a surface, the group $H_{\ell^\infty} := \left\{ x \in \mathrm{GL}_2\left(\mathcal{O}_E \otimes \mathbb{Z}_\ell\right) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}$ appearing in this statement admits a concrete description as follows. When ℓ is split in E, $\mathcal{O}_E \otimes \mathbb{Z}_\ell$ is isomorphic to $\mathbb{Z}_\ell \oplus \mathbb{Z}_\ell$, and $H_{\ell^\infty} = \left\{ (h_1, h_2) \in \mathrm{GL}_2(\mathbb{Z}_\ell)^2 \;\middle|\; \det h_1 = \det h_2 \right\}$. If, on the other hand, ℓ is inert in E, then $\mathcal{O}_E \otimes \mathbb{Z}_\ell$ is a domain that contains a canonical copy of $\mathbb{Z}_\ell$ (namely $\mathbb{Z} \otimes \mathbb{Z}_\ell$), and we have $H_{\ell^\infty} = \left\{ x \in \mathrm{GL}_2(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \;\middle|\; \det x \in \mathbb{Z}_\ell^\times \right\}$, where now det is the usual determinant (since $\mathcal{O}_E \otimes \mathbb{Z}_\ell$ is a domain). More generally, if A is of dimension g, then

$$H_{\ell^\infty} = \left\{ (x_\lambda) \in \prod_{\lambda \mid \ell} \mathrm{GL}_2(\mathcal{O}_\lambda) \;\middle|\; \det x_{\lambda_1} = \det x_{\lambda_2} \in \mathbb{Z}_\ell^\times \quad \forall \lambda_1, \lambda_2 \mid \ell \right\},$$

where the product runs over the places of E dividing ℓ.

Theorem 3.1.7. *Suppose R is an order in an indefinite quaternion division algebra and let Δ be the discriminant of R. Suppose furthermore that all endomorphisms of A are already defined over K. If ℓ is larger than $b(2[K:\mathbb{Q}], 4, 2h(A))^{1/2}$, does not divide Δ, and does not ramify in K, then $G_{\ell^\infty} = (R \otimes \mathbb{Z}_\ell)^\times$.*

Remark 3.1.8. Note that in the case of real and quaternionic multiplication we demand that the endomorphisms of A be defined over K, but this is not a severe restriction. Indeed, this condition can be achieved by passing to a finite extension K' of K, and when A is a surface it is known that K'/K can be taken to be of degree at most 2 for the case of real multiplication ([129, Proposition 4.3]), and at most 12 for the quaternionic case ([25, Prop. 2.1]); more complicated (but still explicit) bounds on the degree K'/K are also available in case A is of GL_2-type, cf. again [129].
Replacing K with K' corresponds to killing the group of connected components of G_{ℓ^∞}, i.e. to demanding that the image of Galois be connected. Analogous results (with slightly different bounds) could be stated without this assumption, at the cost of replacing G_{ℓ^∞} by its identity component in the conclusion.

Before moving to the proofs of the three main statements a few more comments are in order. Consider first the hypothesis that ℓ does not lie below any place of bad reduction of A: without any assumption on A, this condition cannot be turned into an inequality only involving $h(A)$. Indeed, the set of primes dividing the places of bad reduction of A/K is not stable under extensions of

scalars, so it cannot be controlled just in terms of the *stable* Faltings height: this is really an arithmetical condition that is hard to avoid. On the plus side, the primes which fail to meet this restriction are often easy to determine in practice, especially when A is explicitly given as the Jacobian of a genus 2 curve.
Also note that in many intermediate lemmas we give estimates in terms of the best possible isogeny bound (cf. section 3.2.2), thus avoiding to use the specific form of the function $b(A/K)$. However, in order to make the final results more readable, we have chosen to express them in a form that only involves the function b; this also has the merit of giving completely explicit bounds.

Let us also briefly review previous work in the area. As already mentioned, Serre [118] proved that for a large class of abelian varieties (that includes surfaces with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$) there exists a number $\ell_1(A, K)$ such that $G_{\ell^\infty} = \mathrm{GSp}_{2\dim A}(\mathbb{Z}_\ell)$ for every ℓ larger than $\ell_1(A, K)$; his result, however, is not effective, in the sense that the proof does not give any bound on $\ell_1(A, K)$. Similarly, Ribet proved in [109] an open image result for abelian varieties of GL_2-type that includes surfaces with real multiplication as a particular case, but that is again non-effective. The case of quaternionic multiplication was treated independently in [93] and in [45] by extending the techniques Serre used to prove his celebrated open image theorem for elliptic curves in [116], but once again these results were not effective. Finally, it is only fair to also mention the results of Dieulefait who, in [24], gives sufficient conditions for the equality $G_{\ell^\infty} = \mathrm{GSp}_4(\mathbb{Z}_\ell)$ to hold at a prime ℓ; the form of these conditions, however, is again such that they do not yield a bound for the largest prime for which the equality $G_\ell = \mathrm{GSp}_4(\mathbb{Z}_\ell)$ fails to hold. The treatment we give of case 7 of theorem 3.3.2, on the other hand, has been inspired by Dieulefait's paper.

To conclude this introduction let us give a brief overview of the organisation of the chapter and of the proof methods. Theorems 3.1.2, 3.1.4, and 3.1.7 will be shown in sections 3.3, 3.4, and 3.5 respectively.
The main input for the proof in the case of trivial endomorphisms ring comes from group theory, complemented by an application of some nontrivial results of Raynaud. After reducing the problem to that of showing the equality $G_\ell = \mathrm{GSp}_4(\mathbb{F}_\ell)$ for ℓ large enough, we recall the classification of the maximal proper subgroups of $\mathrm{GSp}_4(\mathbb{F}_\ell)$ and proceed to show that each of them cannot occur as the image of the Galois representation on $A[\ell]$, at least for ℓ large enough. In most cases, this follows from the so-called isogeny theorem of Masser and Wüstholz [70] [72] (theorem 3.2.2 below): if the residual representation G_ℓ is small, then the Galois module $A[\ell]$ (or $A[\ell] \times A[\ell]$) is nonsimple, a fact that gives rise to isogenies of high degree, eventually contradicting the isogeny theorem for ℓ large enough. In some exceptional cases, however, the representations $A[\ell]$ and $A[\ell] \times A[\ell]$ can be irreducible even if G_ℓ is comparatively very small, and it is to exclude this possibility that we need to invoke Raynaud's results.

For the case of real multiplication our method is quite different from the one of [109]: by appealing more to group theory, we can completely avoid appealing to Chebotarev's theorem, which would be the main obstacle in making Ribet's method effective.

Finally, a general philosophy suggests that – at the level of Galois representations – an abelian variety of dimension $2g$ with quaternionic multiplication by an algebra with center L should behave as an abelian variety of dimension g admitting multiplication by L, and indeed the case of section 3.5 turns

out to be the easiest, the argument being very similar to that for elliptic curves without complex multiplication. More precisely, the Tate module decomposes as two copies of a 2-dimensional Galois representation, and we can apply techniques that are an essentially straightforward generalisation of those employed to show analogous results for elliptic curves, and that go back to Serre [116] (cf. also [71]).

Finally, in appendix 3.6 we show how to bound the index of $\mathrm{End}_{\overline{K}}(A)$ in any order in which it is contained, a result that is needed in the course of the proof of theorem 3.1.4.

3.2 Preliminaries

We collect in this section a number of results that are essentially well-known and that will form the basis for all our further discussion. Specifically, we recall a few fundamental properties of Galois representations attached to abelian varieties and an explicit form (due to Gaudron and Rémond) of the so-called Isogeny Theorem, first proved by Masser and Wüstholz in a seminal series of papers, cf. especially [70] and [72].

3.2.1 Weil pairing, the multiplier of the Galois action

Recall that the choice of a polarization on A equips the Tate module $T_\ell(A)$ with the Weil pairing, a skew-symmetric, Galois-equivariant form

$$\langle \cdot, \cdot \rangle : T_\ell(A) \times T_\ell(A) \to \mathbb{Z}_\ell(1),$$

where $\mathbb{Z}_\ell(1)$ is the 1-dimensional Galois module the action on which is given by the cyclotomic character $\chi_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathbb{Z}_\ell^\times$. The Weil pairing is known to be nondegenerate on $A[\ell]$ as soon as ℓ does not divide the degree of any given K-polarization of A. Note now that the degree of a minimal K-polarization on A is at most $b(A/K)$ by [28, Théorème 1.1]: since all the bounds given in the main theorem are strictly larger than this number, for the proof of this theorem we can restrict ourselves to only working with primes that *do not* divide the degree of a minimal polarization, and for which the Weil pairing is nondegenerate. We will therefore work under the following

Assumption. For all the primes ℓ we work with, the Weil pairing is nondegenerate on $A[\ell]$.

The fact that $\langle \cdot, \cdot \rangle$ is Galois-equivariant means that G_{ℓ^∞} is a subgroup of $\mathrm{GSp}(T_\ell(A), \langle \cdot, \cdot \rangle)$, the group of symplectic similitudes of $T_\ell(A)$ with respect to the Weil pairing, which we will also simply denote $\mathrm{GSp}(T_\ell(A))$. After a choice of basis we can then consider G_{ℓ^∞} (resp. G_ℓ) as being a subgroup of $\mathrm{GSp}_4(\mathbb{Z}_\ell)$ (resp. $\mathrm{GSp}_4(\mathbb{F}_\ell)$).

Our main interest in the Weil pairing comes from its relationship with the determinant (or, more precisely, the multiplier) of the Galois action. Let us describe the connection. The algebraic group GSp_4 is not simple, a fact which often makes it much easier to work with its normal subgroup Sp_4 instead. To describe the mutual relationship between these groups, note that in general, if $\langle \cdot, \cdot \rangle$ is a skew-symmetric form, the **multiplier** of a symplectic similitude A is the only scalar $\nu(A)$ such that $\langle Av, Aw \rangle = \nu(A)\langle v, w \rangle$ for every v, w. The association $A \mapsto \nu(A)$ is then a homomorphism,

whose kernel is the group $\mathrm{Sp}(\langle\cdot,\cdot\rangle)$ of symplectic isometries. In the case of $\mathrm{Gal}(\overline{K}/K)$ we have an exact sequence

$$1 \to \mathrm{Sp}(T_\ell(A)) \to \mathrm{GSp}(T_\ell(A)) \xrightarrow{\nu} \mathbb{Z}_\ell^\times \to \mathbb{Z}_\ell^\times/\chi_\ell(\mathrm{Gal}(\overline{K}/K)) \to 1,$$

so that in order to prove that G_{ℓ^∞} is all of $\mathrm{GSp}(T_\ell(A))$ it suffices to prove that G_{ℓ^∞} contains $\mathrm{Sp}(T_\ell(A))$ and that $G_{\ell^\infty} \xrightarrow{\nu} \mathbb{Z}_\ell^\times$ is surjective, i.e. that $\mathrm{Gal}(\overline{K}/K) \xrightarrow{\chi_\ell} \mathbb{Z}_\ell^\times$ is surjective. This last condition is very easy to check, and in fact the following lemma is all we will need to pass from results on $\mathrm{Sp}(T_\ell(A))$ to results on $\mathrm{GSp}(T_\ell(A))$.

Lemma 3.2.1. *Suppose ℓ does not ramify in K. Then $\mathrm{Gal}(\overline{K}/K) \xrightarrow{\chi_\ell} \mathbb{Z}_\ell^\times$ is surjective.*
In particular, if G_ℓ (resp. G_{ℓ^∞}) contains $\mathrm{Sp}(A[\ell])$ (resp. $\mathrm{Sp}(T_\ell(A))$) and ℓ does not divide the discriminant of K, the equality $G_\ell = \mathrm{GSp}(A[\ell])$ (resp. $G_{\ell^\infty} = \mathrm{GSp}(T_\ell(A))$) holds.

Proof. The claim is equivalent to the fact that for all $n \geq 1$ the equality $[K(\mu_{\ell^n}) : K] = \varphi(\ell^n)$ holds. It suffices to show that $K/\mathbb{Q}$ and $\mathbb{Q}(\mu_{\ell^n})/\mathbb{Q}$ are linearly disjoint, and since the latter is Galois it suffices to show that they intersect trivially. Now $L := K \cap \mathbb{Q}(\mu_{\ell^n})$ is a subfield of K, so ℓ is unramified in L, and is a subfield of $\mathbb{Q}(\mu_{\ell^n})$, so every prime different from ℓ is also unramified in L. It follows that L is unramified everywhere, that is, $L = \mathbb{Q}$ as claimed. The second statement is immediate. □

3.2.2 The isogeny theorem

For future reference we introduce here the main tool that will make all the explicit estimates possible. The crucial result is the isogeny theorem of Masser and Wüstholz [70] [72], in the following completely explicit form proved by Gaudron and Rémond :

Theorem 3.2.2. *(Isogeny Theorem, [28, Theorem 1.4]) Let $b(A/K)$ be as in definition 3.1.1. For every abelian variety A^* defined over K that is K-isogenous to A, there exists a K-isogeny $A^* \to A$ whose degree is bounded by $b(A/K)$.*

It is very likely that the function $b(A/K)$ of definition 3.1.1 is not the best possible one. Let us then introduce another function $b_0(A/K)$, which is by definition the best possible isogeny bound:

Definition 3.2.3. For A/K an abelian variety, let $b_0(A/K)$ be the smallest natural number such that, for every other abelian variety B/K that is K-isogenous to A, there exists a K-isogeny $B \to A$ of degree at most $b_0(A/K)$. Also let $b_0(A/K;d) = \max_{[K':K]\leq d} b_0(A/K')$, where the maximum is taken over the finite extensions of K of degree at most d.

It is clear that the isogeny theorem implies that $b_0(A/K)$ and $b_0(A/K;d)$ are finite, and that $b_0(A/K;d) \leq b(d[K:\mathbb{Q}], \dim A, h(A))$. Whenever possible, we will state our results in terms of b_0 instead of b. In some situations, however, in order to avoid cumbersome expressions involving maxima we simply give bounds in terms of the function b.

3.3 Type I – Trivial endomorphisms

In this section we establish the surjectivity result under the assumption $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. The material is organized as follows: the first paragraph deals with classical results on the structure of subgroups of $\mathrm{GSp}_4(\mathbb{F}_\ell)$, while in the second we collect information on the action of inertia that will allow us to conclude that some exceptional subgroups of $\mathrm{GSp}_4(\mathbb{F}_\ell)$ cannot arise as images of Galois representations. Theorem 3.1.2 easily follows, as shown in the last paragraph.

3.3.1 Group theory for $\mathrm{GSp}_4(\mathbb{F}_\ell)$

We start by recalling a classical result describing subgroups of $\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$ in terms of their action on $\mathbb{P}_3(\mathbb{F}_\ell)$. We will need a few definitions from classical projective geometry (cf. [41], p.7):

Definition 3.3.1. A hyperbolic (resp. elliptic) congruence is the set of all lines in $\mathbb{P}_3(\mathbb{F}_\ell)$ that meet two given skew lines, each defined over $\mathbb{F}_\ell$ (resp. two conjugate lines defined over $\mathbb{F}_{\ell^2}$ but not over $\mathbb{F}_\ell$). We call these two lines the axes of the congruence.
A parabolic congruence is the set of all lines tangent to a non-degenerate ruled quadric along one of its rulings, forming a one-parameter family of flat pencils sharing one line, namely the ruling they are tangent to. We call this line the axis of the congruence.

Mitchell proved in [78] the following classification (see also King's article in [141] for a more modern account of the result):

Theorem 3.3.2. *Let $\ell > 7$. Every maximal proper subgroup G of $\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$ is of one of the following types:*

1. *G stabilizes a point and a plane in $\mathbb{P}_3(\mathbb{F}_\ell)$;*
2. *G stabilizes a parabolic congruence;*
3. *G stabilizes a hyperbolic congruence;*
4. *G stabilizes an elliptic congruence;*
5. *G stabilizes a quadric and has a subgroup of index 2 isomorphic to $\mathrm{GL}_2(\mathbb{F}_\ell)$;*
6. *G stabilizes a quadric and has a subgroup of index 2 isomorphic to $\mathrm{GU}_2(\mathbb{F}_\ell)$;*
7. *G stabilizes a twisted cubic;*
8. *G has order at most 1920.*

The following simple lemma partially reduces the study of $\mathrm{GSp}_4(\mathbb{F}_\ell)$ to the previous classification:

Lemma 3.3.3. *Let $\ell \geq 3$ and G be a subgroup of $\mathrm{GSp}_4(\mathbb{F}_\ell)$ whose projective image $\mathbb{P}G$ contains $\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$. Then G contains $\mathrm{Sp}_4(\mathbb{F}_\ell)$.*

Proof. Consider the kernel $G^{\nu=1}$ of $G \xrightarrow{\nu} \mathbb{F}_\ell^\times$, where ν denotes the multiplier of a symplectic similitude (cf. section 3.2.1). Suppose $G^{\nu=1}$ is a proper subgroup of $\mathrm{Sp}_4(\mathbb{F}_\ell)$: then by [141, Theorem 2.8 on p. 36] the index $[G : G^{\nu=1}]$ is at least ℓ^3. But if this is the case we have

$$|G| = |G^{\nu=1}| \cdot |\nu(G)| \leq \frac{|\mathrm{Sp}_4(\mathbb{F}_\ell)|}{\ell^3} \cdot (\ell - 1) < |\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)|,$$

contradiction. □

It will prove useful to collate Mitchell's classification with the description of the maximal subgroups of $\mathrm{Sp}_4(\mathbb{F}_\ell)$ given in the spirit of Aschbacher's theorem (cf. for example Tables 8.12 and 8.13 of [19]). Among the geometrical classes introduced by Aschbacher we will only need to deal with $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{C}_3$. Recall that a subgroup $\tilde{G}$ of $\mathrm{Sp}_4(\mathbb{F}_\ell)$ is said to be of class

- $\mathcal{C}_1$, if it stabilizes a totally singular or non-singular subspace;
- $\mathcal{C}_2$, if it stabilizes a direct sum decomposition of $\mathbb{F}_\ell^4$ in subspaces of the same dimension;
- $\mathcal{C}_3$, if there exist a prime r and a subgroup of index r of $\tilde{G}$ whose action on $\mathbb{F}_\ell^4$ is $\mathbb{F}_{\ell^r}$-linear for a given $\mathbb{F}_{\ell^r}$-vector space structure on $\mathbb{F}_\ell^4$. More precisely, $\tilde{G}$ is contained in

  $$\left\{A \in \mathrm{Sp}_4(\mathbb{F}_\ell) \;\middle|\; \exists \sigma \in \mathrm{Gal}\left(\mathbb{F}_{\ell^r}/\mathbb{F}_\ell\right) : \forall \lambda \in \mathbb{F}_{\ell^r}, \forall v \in \mathbb{F}_\ell^4 \quad A(\lambda v) = \sigma(\lambda) A v\right\},$$

  and contains as a subgroup of index at most r the set

  $$\left\{A \in \tilde{G} \;\middle|\; \forall \lambda \in \mathbb{F}_{\ell^r}, \forall v \in \mathbb{F}_\ell^4 \quad A(\lambda v) = \lambda A v\right\}.$$

Let us consider what the various G's in Mitchell's list correspond to in an Aschbacher-type classification. Take $\tilde{G}$ to be the maximal subgroup of $\mathrm{Sp}(4, \mathbb{F}_\ell)$ that lifts G. That the following correspondence is indeed correct follows at once by comparing the indices of the various subgroups in Aschbacher's and Mitchell's classification. Let us disregard case (8), which does not have much geometrical content.

- Cases 1 and 2 correspond to maximal parabolic groups stabilizing totally singular subspaces of dimension 1 (the projective point) and 2 (the projective axis of the congruence) respectively, so that $\tilde{G}$ is of class $\mathcal{C}_1$. For case 2, note that every projectivity sends flat pencils to flat pencils and intersections to intersections, so the axis of the congruence (which is the intersection of all the pencils in the congruence) is sent to itself.
- Case 3 corresponds to a group of class $\mathcal{C}_2$ in Aschbacher's classification. The same argument as with the parabolic congruence shows that every element of $\tilde{G}$ either fixes the axes of the congruence or it interchanges them. Let H be the index-2 subgroup of those $\gamma \in \tilde{G}$ that fix both axes. These axes correspond to trivially-intersecting planes Π_1, Π_2 in $\mathbb{F}_\ell^4$, and $\tilde{G}$ is contained in the stabilizer of the direct sum decomposition $\Pi_1 \oplus \Pi_2$. The group H is isomorphic to $\mathrm{SL}_2(\mathbb{F}_\ell) \times \mathrm{SL}_2(\mathbb{F}_\ell)$, where the two factors act separately on the two planes.
- Case 4 concerns groups belonging to class $\mathcal{C}_3$, with $r = 2$: they admit a subgroup of index 2 isomorphic to $\mathrm{SL}_2(\mathbb{F}_{\ell^2})$, acting naturally on $\mathbb{F}_\ell^4 \cong (\mathbb{F}_{\ell^2})^2$.

- Case 5 corresponds again to class $\mathcal{C}_2$. This is most easily seen by giving an explicit realization of the index-2 subgroup H of $\tilde{G}$: up to conjugation, we can take H to be
$$\left\{ \begin{pmatrix} A & 0 \\ 0 & (A^{-1})^t \end{pmatrix} \mid A \in \mathrm{GL}_2(\mathbb{F}_\ell) \right\},$$
so that (as it is immediate to check) H preserves the symmetric quadratic form whose matrix is $Q = \begin{pmatrix} 0 & \mathrm{Id}_2 \\ \mathrm{Id}_2 & 0 \end{pmatrix}$. Note also that H is symplectic with respect to the standard form $\begin{pmatrix} 0 & -\mathrm{Id}_2 \\ \mathrm{Id}_2 & 0 \end{pmatrix}$. From this description it is immediate to see that $\tilde{G}$ is contained in the stabilizer of the decomposition of $\mathbb{F}_\ell^4$ as direct sum of the two planes defined by the first (resp. last) two coordinates being 0.

- In case (6), $\tilde{G}$ has a subgroup of index 2 that is compatible with an action of $\mathbb{F}_{\ell^2}$ on $\mathbb{F}_\ell^4$, so by definition it belongs to class $\mathcal{C}_3$.

- Finally, groups pertaining to case (7) belong to the exceptional class $\mathcal{S}$ in Aschbacher's classification. By [19], Table 8.13, such $\tilde{G}$'s are isomorphic to $\mathrm{SL}_2(\mathbb{F}_\ell)$.

From this analysis we deduce:

Lemma 3.3.4. *Let $\tilde{G}$ be a subgroup of $\mathrm{GSp}(4, \mathbb{F}_\ell)$ such that $G := \mathbb{P}\tilde{G}$ is contained in $\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$.*

- *If G is contained in a group of type (1), (2), (3) or (5) of Mitchell's list, then $\tilde{G}$ admits a subgroup $\tilde{H}$ of index at most 2 whose action on $\mathbb{F}_\ell^4$ is not irreducible.*

- *If G is contained in a group of type (4) or (6) of Mitchell's list, then $\tilde{G}$ admits a subgroup $\tilde{H}$ of index at most 2 whose action on $\mathbb{F}_\ell^4$ commutes with $\mathbb{F}_{\ell^2}$ (for a suitable structure of $\mathbb{F}_\ell^4$ as $\mathbb{F}_{\ell^2}$-vector space).*

- *If G is contained in a group of type (8) then the subgroup of homotheties of $\tilde{G}$ has index at most* 1920 *(in $\tilde{G}$).*

Proof. We just need to reduce the case of $\mathrm{GSp}_4(\mathbb{F}_\ell)$ to that of $\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$. Denote $\pi : \tilde{G} \to G$ the quotient map. If G falls into case (1), (2), (3) or (5), then it admits a subgroup H of index at most 2 that fixes a point or a line in $\mathbb{P}_3(\mathbb{F}_\ell)$. This point (line) corresponds to a line (plane) in $\mathbb{F}_\ell^4$ that is fixed by any matrix in $\mathrm{GL}_4(\mathbb{F}_\ell)$ lifting an element of H. In particular this is true for every element in the group $\tilde{H} = \pi^{-1}(H)$, which has index at most 2 in $\tilde{G}$ (note that π is surjective, so $[\pi^{-1}(G) : \pi^{-1}(H)] = [G : H]$).
Next suppose G is contained in a group of type (4) or (6). Let $\pi_1 : \mathrm{Sp}_4(\mathbb{F}_\ell) \to \mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$ be the canonical projection and $G_1 = \pi_1^{-1}(G)$. The hypothesis $G \subseteq \mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$ implies that $\pi_1(G_1) = \pi(\tilde{G})$: every element of G lifts to an element of $\mathrm{Sp}_4(\mathbb{F}_\ell)$. The group G_1 contains a subgroup H_1 of index at most 2 that commutes with the action of $\mathbb{F}_{\ell^2}$. Define $\tilde{H} = \pi^{-1}(\pi_1(H_1))$. As before,
$$[\tilde{G} : \tilde{H}] = [\pi(\tilde{G}) : \pi(\tilde{H})] = [\pi_1(G_1) : \pi_1(H_1)] = [G_1 : H_1] \leq 2,$$
and furthermore an element in $\tilde{H}$ differs from an element in H_1 by a homothety, so $\tilde{H}$ commutes with $\mathbb{F}_{\ell^2}$ since this is true for H_1.

Finally, if G is of type (8) then the trivial group has index at most 1920 in $\tilde{G}$, so $\pi^{-1}(\mathrm{Id})$ (which consists entirely of homotheties) has index at most 1920 in G. □

To conclude this group-theoretic part we describe in some more detail the relationship between subgroups of $\mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$ and $\mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$ and the structure of groups falling into case (7). We shall need the following fact, which is well-known and can be found for example in [19] (Tables 8.13 and 8.14, column "Stabilizer"):

Lemma 3.3.5. *Every maximal subgroup of $\mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$ not containing $\mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$ is an extension of degree at most 2 of a proper maximal subgroup of $\mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$.*

Definition 3.3.6. We shall say that a maximal subgroup G of $\mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$ is of type n, with $1 \leq n \leq 8$, if G is an extension (of degree at most 2) of a maximal subgroup of $\mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$ of the corresponding type.

For groups of type (7) we have the following precise description ([19, §5.3], especially Proposition 5.3.6); notice that the condition $\ell \geq 11$ ensures the existence of maximal subgroups of $\mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$ of type (7), cf. [19, Table 8.13].

Lemma 3.3.7. *Let $\ell \geq 11$ be a prime, $G_0 \subseteq \mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$ be a maximal subgroup of type (7) and $G \subseteq \mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$ a maximal subgroup of $\mathbb{P}\mathrm{GSp}_4(\mathbb{F}_\ell)$ containing G_0 as an index-2 subgroup. Then G is isomorphic to $\mathrm{PGL}_2(\mathbb{F}_\ell)$, and for every $g \in G$ and for every $\gamma \in \mathrm{GSp}_4(\mathbb{F}_\ell)$ lifting g, the eigenvalues of γ can be written as $\mu\lambda_1^3, \mu\lambda_1^2\lambda_2, \mu\lambda_1\lambda_2^2, \mu\lambda_2^3$, where λ_1 and λ_2 are the roots of a second-degree polynomial with coefficients in $\mathbb{F}_\ell$ and μ is an element of $\mathbb{F}_\ell^\times$.*

Remark 3.3.8. This can also be deduced from completely abstract considerations: the unique irreducible 4-dimensional representation of the algebraic group SL_2 is symplectic, so it gives an embedding $\mathrm{SL}_2 \hookrightarrow \mathrm{Sp}_4$ with weights $-3, -1, 1, 3$. This representation extends to a map $\mathrm{PGL}_2 \to \mathbb{P}\mathrm{GSp}_4$: it is then not hard to see that this situation must correspond to case (7) above. The λ_i's, $i = 1, 2$, are the eigenvalues of the 2 by 2 matrix corresponding to γ in GL_2.

3.3.2 The action of inertia

In this section $\mathfrak{l}$ denotes a prime of $\mathcal{O}_K$ of good reduction for A and $I_\mathfrak{l}$ the *tame* inertia group at $\mathfrak{l}$. Let ℓ be the rational prime below $\mathfrak{l}$ and e the absolute ramification index of $\mathfrak{l}$. We recall the following well-known result of Raynaud:

Theorem 3.3.9. *([104, Corollaire 3.4.4]) Let V be a simple Jordan-Hölder quotient of $A[\ell]$ (as a module over the inertia group at $\mathfrak{l}$). Suppose that V has dimension n over $\mathbb{F}_\ell$. The action of the inertia group at $\mathfrak{l}$ on $A[\ell]$ factors through $I_\mathfrak{l}$. Moreover, there exist integers $e_1, \ldots, e_n$ such that:*

- *V has a structure of $\mathbb{F}_{\ell^n}$-vector space*
- *the action of $I_\mathfrak{l}$ on V is given by a character $\psi : I_\mathfrak{l} \to \mathbb{F}_{\ell^n}^\times$*
- *$\psi = \varphi_1^{e_1} \ldots \varphi_n^{e_n}$, where $\varphi_1, \ldots, \varphi_n$ are the fundamental characters of $I_\mathfrak{l}$ of level n*

- *for every $i = 1, \dots, n$ the inequality $0 \leq e_i \leq e$ holds*

In this section we only concern ourselves with the action of $I_\mathfrak{l}$ on $A[\ell]$; in particular, by a Jordan-Hölder quotient of $A[\ell]$ we implicitly mean "under the action of $I_\mathfrak{l}$".

Convention. There is of course a certain ambiguity in the numbering of the characters of level n. We choose our numbering so that $\varphi_j = (\varphi_1)^{\ell^{j-1}}$ for $j = 1, \dots, n$. Note that the norm, taken from $\mathbb{F}_{\ell^n}$ to $\mathbb{F}_\ell$, of the character φ_1 (hence of all characters of level n) is the unique fundamental character of level 1. When ℓ is unramified in K, this fundamental character of level 1 is χ_ℓ, the cyclotomic character mod ℓ.

Recall the following fact (a consequence of the definition of the fundamental characters):

Lemma 3.3.10. *If $\varphi : I_\mathfrak{l} \to \mathbb{F}_{\ell^n}^\times$ is any fundamental character of order n, then φ is surjective, hence in particular its image is a cyclic group of order $\ell^n - 1$.*

Corollary 3.3.11. *Suppose ℓ is at least 5 and unramified in K, and let V be an n-dimensional Jordan-Hölder quotient of $A[\ell]$. Then the following are the only possibilities:*

- *$I_\mathfrak{l}$ acts trivially*
- *$I_\mathfrak{l}$ acts through a fundamental character of order n*
- *$I_\mathfrak{l}$ acts through the product of two distinct fundamental characters of order n, where $n \geq 3$*

Proof. For any simple Jordan-Hölder constituent W of $A[\ell]$ denote by $\psi_W = \varphi_1^{e_1(W)} \dots \varphi_n^{e_n(W)}$ the character giving the action of $I_\mathfrak{l}$ on W, where n is the dimension of W. The determinant of the action of $I_\mathfrak{l}$ on W is the norm $N_{\mathbb{F}_{\ell^n}/\mathbb{F}_\ell}\left(\varphi_1^{e_1(W)} \dots \varphi_{n(W)}^{e_n(W)}\right) = \chi_\ell^{e_1(W)+\dots+e_n(W)}$. Since the determinant of the Galois action on $A[\ell]$ is χ_ℓ^2 by the properties of the Weil pairing, we must have

$$\prod_{\substack{W \text{ Jordan-Hölder} \\ \text{factor of } A[\ell]}} \chi_\ell^{e_1(W)+\dots+e_n(W)}(g) = \chi_\ell(g)^2 \quad \forall g \in I_\mathfrak{l},$$

where the product is taken over a fixed Jordan-Hölder filtration of $A[\ell]$ that contains V. Comparing orders we deduce

$$\sum_{\substack{W \text{ Jordan-Hölder} \\ \text{factor of } A[\ell]}} (e_1(W) + \dots + e_n(W)) \equiv 2 \pmod{|\chi_\ell(I_\mathfrak{l})|}. \tag{3.1}$$

Now since every $e_i(W)$ is at most 1 (as ℓ is unramified in K we have $e_i(W) \leq e(\ell) = 1$) the left hand side is at most $\sum_{\substack{W \text{ Jordan-Hölder} \\ \text{factor of } A[\ell]}} \dim(W) = 4$, hence from the inequality $4 \leq \ell - 1 = |\chi_\ell(I_\mathfrak{l})|$ (cf. lemma 3.3.10) we deduce that the congruence of equation (3.1) must in fact be an equality, and

$$\sum_{\substack{W \text{ Jordan-Hölder} \\ \text{factor of } A[\ell]}} (e_1(W) + \dots + e_n(W)) = 2.$$

In particular, taking again into account the fact that $e_i(V) \leq 1$, the only possibilities for the character giving the action of $I_\mathfrak{l}$ on V are precisely those given in the statement. To see the necessity of the condition $n \geq 3$ in the last case, note that for $n = 1$ there are no two distinct fundamental characters and for $n = 2$ the product $\varphi_1\varphi_2$ coincides with χ_ℓ, so the action would factor through $\mathbb{F}_\ell^\times$ and the Jordan-Hölder factor V would not be simple. □

Proposition 3.3.12. *Under the hypotheses of the previous lemma, for every $g \in I_\mathfrak{l}$ the multiset of eigenvalues of $\rho_\ell(g)$ is one of the following (the superscript indicates the level of the characters):*

1. $\left\{\varphi_1^{(4)}(g)\varphi_2^{(4)}(g), \varphi_2^{(4)}(g)\varphi_3^{(4)}(g), \varphi_3^{(4)}(g)\varphi_4^{(4)}(g), \varphi_4^{(4)}(g)\varphi_1^{(4)}(g)\right\}$
2. $\left\{\varphi_1^{(3)}(g)\varphi_2^{(3)}(g), \varphi_2^{(3)}(g)\varphi_3^{(3)}(g), \varphi_3^{(3)}(g)\varphi_1^{(3)}(g), 1\right\}$
3. $\left\{\varphi_1^{(3)}(g), \varphi_2^{(3)}(g), \varphi_3^{(3)}(g), \chi_\ell(g)\right\}$
4. $\left\{\varphi_1^{(2)}(g), \varphi_2^{(2)}(g), \varphi_1^{(2)}(g), \varphi_2^{(2)}(g)\right\}$
5. $\left\{\varphi_1^{(2)}(g), \varphi_2^{(2)}(g), \chi_\ell(g), 1\right\}$
6. $\{\chi_\ell(g), \chi_\ell(g), 1, 1\}$

Proof. The multiset of eigenvalues of $\rho_\ell(g)$ is the union of the multisets of values taken by the characters that give the action on the simple Jordan-Hölder factors of $A[\ell]$. With this remark and the previous lemma at hand the rest of the proof is just casework. The cases in the list correspond to decompositions of $A[\ell]$ with simple Jordan-Hölder factors of dimensions respectively 4 (case 1), 3+1 (cases 2 and 3), 2+2 (case 4), 2+1+1 (case 5) and 1+1+1+1 (case 6).
Note that the multiset $\left\{\varphi_1^{(4)}(g)\varphi_3^{(4)}(g), \varphi_2^{(4)}(g)\varphi_4^{(4)}(g), \varphi_3^{(4)}(g)\varphi_1^{(4)}(g), \varphi_4^{(4)}(g)\varphi_2^{(4)}(g)\right\}$, which does not appear in the above list, is actually the same as $\left\{\varphi_1^{(2)}(g), \varphi_2^{(2)}(g), \varphi_1^{(2)}(g), \varphi_2^{(2)}(g)\right\}$. □

Remark 3.3.13. In cases 1 through 3 the inertia group $I_\mathfrak{l}$ contains at least one g such that the set of eigenvalues of $\rho_\ell(g)$ is contained neither in $\mathbb{F}_\ell^\times$ nor in $\mathbb{F}_{\ell^2}^\times$: otherwise, the action of $I_\mathfrak{l}$ on each Jordan-Hölder quotient of $A[\ell]$ would factor through $\mathbb{F}_{\ell^2}^\times$ and there would be no simple Jordan-Hölder quotient of dimension at least 3.

We deduce the following technical result which will come in handy later:

Proposition 3.3.14. *Suppose ℓ is at least 11 and unramified in K: then $\mathbb{P}G_\ell$ is not contained in a group of type (7).*

Proof. Suppose on the contrary that $\mathbb{P}G_\ell$ is of type (7). By lemma 3.3.7, for any $g \in \mathrm{Gal}(\overline{K}/K)$ the eigenvalues of $\rho_\ell(g)$ are of the form $\{\mu a^3, \mu a^2 d, \mu a d^2, \mu d^3\}$ for certain $a, d \in \mathbb{F}_{\ell^2}^\times$ and $\mu \in \mathbb{F}_\ell^\times$. This applies in particular to the tame inertia group $I_\mathfrak{l}$: for every g in $I_\mathfrak{l}$, the eigenvalues of $\rho_\ell(g)$ lie in $\mathbb{F}_{\ell^2}^\times$, and – taken in some order $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ – they satisfy the system of equations

$$\lambda_1\lambda_4 = \lambda_2\lambda_3,\ \lambda_2\lambda_4 = \lambda_3^2,\ \lambda_1\lambda_3 = \lambda_2^2. \tag{3.2}$$

We will now go through all the cases listed in proposition 3.3.12 and see that (for a suitably chosen $g \in I_\mathfrak{l}$) there is no way to renumber the multiset of eigenvalues of $\rho_\ell(g)$ in such a way that the three equations above are all satisfied together, a contradiction that shows the result. Remark 3.3.13 implies that cases 1 through 3 do not happen (since we have just seen that the eigenvalues of every $\rho_\ell(g)$ are in $\mathbb{F}_{\ell^2}^\times$). Next we consider case 6. Note that the condition $\ell \geq 11$ implies that the order of χ_ℓ is at least 10 (by lemma 3.3.10), so there exists a $g \in I_\mathfrak{l}$ with $\chi_\ell(g) \neq \pm 1$. Consider equations (3.2) for this specific g. If either λ_2 or λ_3 is 1, then one of the last two equations reads $\chi_\ell(g)^d = 1$

with $d = 1$ or 2, which contradicts $\chi_\ell(g) \neq 1, -1$. But if neither λ_2 nor λ_3 is 1 then the only possibility is $\lambda_1 = \lambda_4 = 1$, $\lambda_2 = \lambda_3 = \chi_\ell(g)$, which violates *all* of the three equations.
Likewise, in case 4 we can choose a $g \in I_\mathfrak{l}$ such that $\varphi_1^{(2)}(g)$ is of order $\ell^2 - 1$, and – independently of the numbering of the eigenvalues – from equations (3.2) we obtain $\left(\varphi_1^{(2)}(g)\right)^2 = \left(\varphi_2^{(2)}(g)\right)^2$, which using $\varphi_2^{(2)} = \left(\varphi_1^{(2)}\right)^\ell$ implies $\left(\varphi_1^{(2)}(g)\right)^{2(\ell-1)} = 1$, a contradiction.
Finally, if we are in case 5, then taking the norm from $\mathbb{F}_{\ell^2}$ to $\mathbb{F}_\ell$ of the three equations (3.2) we find that for all $g \in I_\mathfrak{l}$ there exists a positive integer $d \leq 3$ such that $\chi_\ell(g)^d = 1$, which again contradicts the fact that $\chi_\ell(I_\mathfrak{l})$ has order at least 10. □

Remark 3.3.15. In chapter 4 we study in greater generality those maximal subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ whose projective image is isomorphic to either $\mathrm{PSL}_2(\mathbb{F}_\ell)$ or $\mathrm{PGL}_2(\mathbb{F}_\ell)$, deducing results similar to proposition 3.3.14 that hold in arbitrary dimension. The method of chapter 4 also requires much less casework.

3.3.3 The surjectivity result

We are almost ready to prove theorem 3.1.2. The ingredients we are still missing are two isogeny estimates, which form the subject of lemmas 3.3.17 and 3.3.18 below, and the following result due to Serre:

Lemma 3.3.16. *Let n be a positive integer, $\ell \geq 5$ be a prime, and H be a closed subgroup of $\mathrm{Sp}_{2n}(\mathbb{Z}_\ell)$ whose projection modulo ℓ contains $\mathrm{Sp}_{2n}(\mathbb{F}_\ell)$: then $H = \mathrm{Sp}_{2n}(\mathbb{Z}_\ell)$. Likewise, let G be a closed subgroup of $\mathrm{GSp}_{2n}(\mathbb{Z}_\ell)$ whose projection modulo ℓ contains $\mathrm{Sp}_{2n}(\mathbb{F}_\ell)$: then $G' = \mathrm{Sp}_{2n}(\mathbb{Z}_\ell)$.*

Proof. The first statement is [120, Lemme 1 on p. 52]. The second part follows from applying the first to $G = H'$: indeed, the image modulo ℓ of H' contains the derived subgroup of $\mathrm{Sp}_{2n}(\mathbb{F}_\ell)$, which (since $\ell \geq 5$) is again $\mathrm{Sp}_{2n}(\mathbb{F}_\ell)$, and the claim follows. □

Lemma 3.3.17. *Let A/K be an abelian variety of dimension g with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, and let ℓ be a prime strictly larger than $b_0(A \times A/K)^{1/2g}$. The centralizer of G_ℓ inside $\mathrm{End}(A[\ell])$ is $\mathbb{F}_\ell$.*

Proof. Suppose that the centralizer of G_ℓ inside $\mathrm{End}(A[\ell])$ is strictly larger than $\mathbb{F}_\ell$ and choose an α lying in this centralizer but not in $\mathbb{F}_\ell$. Consider the abelian variety $B = A \times A$ and the subgroup of B given by $\Gamma = \{(x, \alpha x) \mid x \in A[\ell]\}$. Note that Γ is defined over K: indeed any $g \in \mathrm{Gal}(\overline{K}/K)$ sends $(x, \alpha x)$ to $(\rho_\ell(g)x, \rho_\ell(g)\alpha x) = (\rho_\ell(g)x, \alpha(\rho_\ell(g)x)) \in \Gamma$ (since α commutes with all of $\rho_\ell(\mathrm{Gal}(\overline{K}/K))$).
Let $B^* = B/\Gamma$, $\pi : B \to B^*$ be the natural projection and $\psi : B^* \to B$ be an isogeny in the opposite direction satisfying $\deg(\psi) \leq b_0(A \times A/K)$. The isogeny $\psi \circ \pi$ of B kills Γ, and on the other hand by the hypothesis $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$ it is representable as a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a, b, c, d \in \mathbb{Z}$. By definition we must have $ax + b\alpha x = cx + d\alpha x = 0$ for every $x \in A[\ell]$. Suppose that one among a, b, c, d is not divisible by ℓ (and for the sake of simplicity let us assume it is b): then $\alpha(x) = -b^{-1}ax$ for every x in $A[\ell]$, which shows that α is multiplication by an element of $\mathbb{F}_\ell$, contradiction. Therefore a, b, c, d are all divisible by ℓ, and the degree of $\psi \circ \pi$, which is $\left(\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)^{2g}$, is divisible by ℓ^{4g}.

On the other hand, the degree of π is $|\Gamma| = |A[\ell]| = \ell^{2g}$, so we deduce $\ell^{2g} \leq b_0(A \times A/K)$, which contradicts the hypothesis. □

Lemma 3.3.18. *Let A/K be an abelian variety of dimension g with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, and let ℓ be a prime strictly larger than $b_0(A/K)$. The G_ℓ-module $A[\ell]$ is irreducible.*

Proof. Let ℓ be such that $A[\ell]$ is not irreducible and let H be a nontrivial subspace of $A[\ell]$ stable under the action of G_ℓ. As H is a proper $\mathbb{F}_\ell$-subspace of $A[\ell] \cong \mathbb{F}_\ell^{2g}$, its order divides ℓ^{2g-1}. Consider now the abelian variety $A^* = A/H$, which is defined over K (since H is); we know that there exists a nontrivial isogeny $\Psi : A^* \to A$ of degree at most $b_0(A/K)$. Let $\pi : A \to A^*$ be the canonical projection, of degree $|H|$ (which divides ℓ^{2g-1}), and consider the composition $\Psi \circ \pi : A \to A$. By the hypothesis $\mathrm{End}(A) = \mathbb{Z}$ this composition must be multiplication by m for a certain nonzero integer m. Comparing degrees we see that $m^{2g} = \deg(\Psi)\deg(\pi) \leq b_0(A/K)\ell^{2g-1}$, and on the other hand $\Psi \circ \pi$ kills H (since this is true even for π alone), so $mH = 0$. Every nonzero element of H has order ℓ, so m must be divisible by ℓ, which implies $\ell^{2g} \leq b_0(A/K)\ell^{2g-1}$, i.e. $\ell \leq b_0(A/K)$. □

Theorem 3.3.19. *Let A/K be an Abelian surface with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. Let ℓ be a rational prime that is not below any place of bad reduction of A. Suppose that ℓ does not ramify in K and is strictly larger than $b(2 \cdot 1920[K : \mathbb{Q}], 4, 2h(A))^{1/4}$: then $G_{\ell^\infty} = \mathrm{GSp}_4(\mathbb{Z}_\ell)$.*

Proof. By lemma 3.2.1 we just need to show that G_{ℓ^∞} contains $\mathrm{Sp}_4(\mathbb{Z}_\ell)$, by lemma 3.3.16 it is enough to prove that G_ℓ contains $\mathrm{Sp}_4(\mathbb{F}_\ell)$, and by lemma 3.3.3 we are reduced to showing that $\mathbb{P}G_\ell$ contains $\mathbb{P}\,\mathrm{Sp}_4(\mathbb{F}_\ell)$. Suppose the contrary: then $\mathbb{P}G_\ell$ is contained in a group of one of the types (1) through (8) of definition 3.3.6; by proposition 3.3.14, type (7) is excluded. Let $K_{\ell,1}$ be the quadratic extension of K defined by

$$\ker\left(\mathrm{Gal}(\overline{K}/K) \to \mathbb{P}G_\ell \to \frac{\mathbb{P}G_\ell}{(\mathbb{P}G_\ell \cap \mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell))}\right),$$

and let $H_\ell := \rho_\ell\left(\mathrm{Gal}(\overline{K_{\ell,1}}/K_{\ell,1})\right)$. By construction, $\mathbb{P}H_\ell$ is a proper subgroup of $\mathbb{P}\mathrm{Sp}_4(\mathbb{F}_\ell)$: in particular, $\mathbb{P}H_\ell$ is contained in one of the groups in Mitchell's list, and we have already excluded case (7). Hence by lemma 3.3.4 there is a subgroup J_ℓ of H_ℓ of index at most 1920 that either admits an invariant subspace in $A[\ell]$ or commutes with an action of $\mathbb{F}_{\ell^2}$, and by Galois theory J_ℓ corresponds to a certain extension $K_{\ell,2}/K_{\ell,1}$ of degree at most 1920. This extension $K_{\ell,2}$ has the property that $\rho_\ell(\mathrm{Gal}(\overline{K_{\ell,2}}/K_{\ell,2})) = J_\ell$ either admits an invariant subspace or commutes with an action of $\mathbb{F}_{\ell^2}$: but this contradicts lemma 3.3.18 or 3.3.17 respectively, where we can safely replace the function b_0 by the function b thanks to theorem 3.2.2. This contradiction establishes the theorem. □

Remark 3.3.20. By the methods of chapter 4 (cf. proposition 4.6.5) we can give a uniform bound on the largest prime for which G_ℓ can be of type (8): this has the effect of improving the bound to $b(4[K : \mathbb{Q}], 4, 2h(A))^{1/4}$, and with some more effort to $b(2[K : \mathbb{Q}], 4, 2h(A))^{1/4}$.

3.4 Type I – Real multiplication

We consider now the case of GL_2-varieties, which includes abelian surfaces with real multiplication as a special case. Recall that an abelian variety A is said to be of GL_2-type when it is absolutely simple and $\mathrm{End}_{\overline{K}}(A)$ is an order in a totally real number field E of degree equal to $\dim A$; we shall assume that this action is defined over K. For every ℓ we put $\mathcal{O}_\ell = \mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$, and if λ is a place of E we let $\mathcal{O}_\lambda$ be the completion of $\mathcal{O}_E$ at λ. We have $\mathcal{O}_\ell \cong \prod_{\lambda|\ell} \mathcal{O}_\lambda$, where the product is over the places of E dividing ℓ. An implicit convention will always be in force, that if λ is a place of E then ℓ denotes its residual characteristic.

Definition 3.4.1. Following Ribet's paper [109] we say that a rational prime ℓ is **good** for A if it does not divide the index $[\mathcal{O}_E : R]$.

As $[\mathcal{O}_E : R]$ is finite, all but finitely many primes are good for A. It is a general fact that $[\mathcal{O}_E : R]$ can be bounded in terms of K and $h(A)$, cf. appendix 3.6. We obtain from proposition 3.6.5 the following fact, which enables us to assume that all the primes we work with are good:

Proposition 3.4.2. *The index $[\mathcal{O}_E : \mathrm{End}(A)]$ is bounded by $b\left([K:\mathbb{Q}], \dim A, h(A)\right)^{\dim A}$. In particular, any ℓ strictly larger than this quantity is good.*

From now on we only consider good primes – this only excludes a finite, explicitly bounded number of cases. Notice that in the case of surfaces, in view of the last proposition and of the obvious inequality $b(2[K:\mathbb{Q}], 2\dim A, 2h(A))^{1/2} > b\left([K:\mathbb{Q}], \dim A, h(A)\right)^2$, all the primes considered in corollary 3.1.5 are good for A. For any good prime ℓ we have $R_\ell := R \otimes \mathbb{Z}_\ell \cong \mathcal{O}_E \otimes \mathbb{Z}_\ell$, and furthermore

Proposition 3.4.3. *([109], Proposition 2.2.1) If ℓ is good for A, then $T_\ell(A)$ is a free R_ℓ-module of rank 2; equivalently, it is a free $\mathcal{O}_\ell$-module of rank 2.*

When ℓ is good and λ is a place of E of characteristic ℓ we put $T_\lambda(A) = T_\ell(A) \otimes_{\mathcal{O}_\ell} \mathcal{O}_\lambda$: this makes sense since $\mathcal{O}_\ell = R \otimes \mathbb{Z}_\ell$. The Galois action on $T_\ell(A)$ is $\mathcal{O}_\ell$-linear, and we thus obtain canonical decompositions $T_\ell(A) \cong \prod_{\lambda|\ell} T_\lambda(A)$; the $\mathcal{O}_\ell$-linear morphism $\rho_{\ell\infty}$ then amounts to a family of $\mathcal{O}_\lambda$-linear maps

$$\rho_{\lambda\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(T_\lambda(A)) \cong \mathrm{GL}_2\left(\mathcal{O}_\lambda\right).$$

We also have isomorphisms $\mathrm{Aut}\, T_\ell(A) \cong \mathrm{GL}_2(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \cong \prod_{\lambda|\ell} \mathrm{GL}_2\left(\mathcal{O}_\lambda\right)$, and we regard the ℓ-adic Galois representation on $T_\ell(A)$ as a group morphism

$$\rho_{\ell\infty} : \mathrm{Gal}(\overline{K}/K) \to \prod_{\lambda|\ell} \mathrm{GL}_2\left(\mathcal{O}_\lambda\right).$$

It is also natural to consider λ-adic residual representations:

Definition 3.4.4. If λ is a place of E above a good prime ℓ we write G_λ for the image of the residual representation modulo λ, namely the image of the map ρ_λ given by the composition

$$\mathrm{Gal}(\overline{K}/K) \overset{\rho_{\ell\infty}}{\to} \prod_{\lambda|\ell} \mathrm{GL}_2\left(\mathcal{O}_\lambda\right) \to \mathrm{GL}_2\left(\mathcal{O}_\lambda\right) \to \mathrm{GL}_2\left(\mathcal{O}_\lambda/\lambda\right).$$

The determinant of $\rho_{\lambda\infty}$ is easy to describe:

Lemma 3.4.5. *([109], Lemma 4.5.1) For every λ dividing a good prime, the function*

$$\det_{\mathcal{O}_\lambda} \rho_{\lambda\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathcal{O}_\lambda^\times$$

agrees with $\chi_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathbb{Z}_\ell^*$, the ℓ-adic cyclotomic character.

Observe that for a good prime ℓ the ℓ-adic representation lands in $\mathrm{Aut}_{\mathcal{O}_\ell}(T_\ell A)$. If we regard $\mathbb{Z}_\ell^*$ as being embedded in $\mathcal{O}_\ell$ by the fact that the latter is naturally a $\mathbb{Z}_\ell$-algebra, the determinant of $\rho_{\ell\infty}(g)$ with respect to $\mathcal{O}_\ell$ is an element of $\mathbb{Z}_\ell^*$, and the previous result (combined with lemma 3.2.1) gives

Lemma 3.4.6. *If ℓ is good and unramified in K then $\det_{\mathcal{O}_\ell} : G_{\ell\infty} \to \mathbb{Z}_\ell^*$ is surjective.*

3.4.1 The intersection $G_{\ell\infty} \cap \mathrm{SL}_2(\mathcal{O}_\ell)$

The key step in proving the surjectivity of the Galois representation for ℓ large enough lies in understanding the intersection $G_{\ell\infty} \cap \mathrm{SL}_2(\mathcal{O}_\ell)$. A remarkable simplification of the problem comes from the fact that we can limit ourselves to studying the residual mod-ℓ representation instead of the full ℓ-adic system: this is made possible by the following 'lifting' result, analogous to lemma 3.3.16.

Proposition 3.4.7. *([113], Proposition 4.2) Let $\mathcal{O}$ be the ring of integers of a number field E, $\lambda_1, \lambda_2, \ldots, \lambda_r$ distinct primes of $\mathcal{O}$ above ℓ and H a closed subgroup of $\mathrm{SL}_2(\mathcal{O}_{\lambda_1}) \times \cdots \times \mathrm{SL}_2(\mathcal{O}_{\lambda_r})$ whose projection to $\mathrm{SL}_2(\mathbb{F}_{\lambda_1}) \times \cdots \times \mathrm{SL}_2(\mathbb{F}_{\lambda_r})$ is surjective. If ℓ is unramified in E and $\ell \geq 5$, then H is all of $\mathrm{SL}_2(\mathcal{O}_{\lambda_1}) \times \cdots \times \mathrm{SL}_2(\mathcal{O}_{\lambda_r})$. Under the same assumptions on ℓ, if G is a closed subgroup of $\mathrm{GL}_2(\mathcal{O}_{\lambda_1}) \times \cdots \times \mathrm{GL}_2(\mathcal{O}_{\lambda_r})$ whose projection to $\mathrm{GL}_2(\mathbb{F}_{\lambda_1}) \times \cdots \times \mathrm{GL}_2(\mathbb{F}_{\lambda_r})$ contains $\mathrm{SL}_2(\mathbb{F}_{\lambda_1}) \times \cdots \times \mathrm{SL}_2(\mathbb{F}_{\lambda_r})$, then $G' = \mathrm{SL}_2(\mathcal{O}_{\lambda_1}) \times \cdots \times \mathrm{SL}_2(\mathcal{O}_{\lambda_r})$.*

Recall that we only work with good primes: concretely, this means that all the statements to follow have the implicit hypothesis that ℓ is good for A.

3.4.1.1 A little group theory

We briefly review some group-theoretic results we are going to use. One is the following sufficient criterion for a group to be a direct product:

Lemma 3.4.8. *([109], Lemma 5.2.2) Let $S_1, \ldots, S_k$ ($k > 1$) be finite groups with no nontrivial abelian quotients. Let G be a subgroup of $S_1 \times \cdots \times S_k$ such that each projection $G \to S_i \times S_j$ ($1 \leq i < j \leq k$) is surjective. Then $G = S_1 \times \cdots \times S_k$.*

We will also need the following version of [71, Lemma 5.1]; note that even though our hypotheses are slightly different from those of [71] the same proof works in our setting as well.

Lemma 3.4.9. *Let $\ell \geq 5$ be a prime, $\mathbb{F}$ be a finite field of characteristic ℓ, and*

$$D = \left\{(b, b') \in \mathrm{GL}_2(\mathbb{F}) \times \mathrm{GL}_2(\mathbb{F}) \;\middle|\; \det(b) = \det(b') \in \mathbb{F}_\ell^\times\right\}.$$

Let H be a subgroup of D whose projections on the two factors $\mathrm{GL}_2(\mathbb{F})$ contain $\mathrm{SL}_2(\mathbb{F})$. Then either H contains $\mathrm{SL}_2(\mathbb{F}) \times \mathrm{SL}_2(\mathbb{F})$, or there exist an isomorphism $f : V \to V'$, a character $\chi : H \to \{\pm 1\}$ and a (field) automorphism σ of $\mathbb{F}$ such that

$$H \subseteq \left\{(b, b') \in \mathrm{GL}(V) \times \mathrm{GL}(V') \;\middle|\; b' = \chi((b, b'))\sigma\left(fbf^{-1}\right)\right\}.$$

Finally we will need a description of the subgroups of $\mathrm{GL}_2(\mathbb{F}_{\ell^\beta})$ for $\beta \geq 1$:

Theorem 3.4.10. *(Dickson, [12, Theorem 3.4]) Let p be a prime number, β a positive integer, $q = p^\beta$, and G a subgroup of $\mathrm{GL}_2(\mathbb{F}_q)$. Then, up to conjugacy in $\mathrm{GL}_2(\mathbb{F}_q)$, one of the following occurs:*

1. *G is cyclic;*
2. *G is a subgroup of the Borel group $\left\{\begin{pmatrix} x & y \\ 0 & z \end{pmatrix} \;\middle|\; x, z \in \mathbb{F}_q^\times, y \in \mathbb{F}_q\right\}$;*
3. *G contains (as a subgroup of index 2) a cyclic subgroup of order u, where u divides $q^2 - 1$;*
4. *G contains (as a subgroup of index 2) a subgroup consisting entirely of diagonal matrices;*

5a. *$p^\beta > 3$, and there is an $\alpha \in \mathbb{N}_{>0}$ dividing β such that G is generated by $\mathrm{SL}_2(\mathbb{F}_{p^\alpha})$ and by a scalar matrix V;*

5b. *$p^\beta > 3$, and there exist an α dividing β, a generator ε of $\mathbb{F}_{p^\beta}^\times$ (as a multiplicative group), and an element $b \in \mathbb{F}_{p^\beta}^\times$, such that G is generated by $\mathrm{SL}_2(\mathbb{F}_{p^\alpha})$, a scalar matrix V, and the diagonal matrix $\mathrm{diag}(b, b\varepsilon)$; the subgroup generated by $\mathrm{SL}_2(\mathbb{F}_{p^\alpha})$ and V is of type 5a, and has index 2 in G;*

6. *$G/\{\pm \mathrm{Id}\}$ is isomorphic to $S_4 \times \frac{\mathbb{Z}}{u\mathbb{Z}}$, $A_4 \times \frac{\mathbb{Z}}{u\mathbb{Z}}$ or $S_5 \times \frac{\mathbb{Z}}{u\mathbb{Z}}$, where $\frac{\mathbb{Z}}{u\mathbb{Z}}$ is identified with the subgroup generated by a scalar matrix in $\mathrm{GL}_2(\mathbb{F}_q)/\{\pm \mathrm{Id}\}$.*
7. *G is not of type (6), but $G/\{\pm \mathrm{Id}\}$ contains $A_4 \times \frac{\mathbb{Z}}{u\mathbb{Z}}$ as a subgroup of index 2, and A_4 as a subgroup with cyclic quotient group; $\frac{\mathbb{Z}}{u\mathbb{Z}}$ is as in type (6) with u even.*

Definition 3.4.11. In cases (5a) or (5b) the number α will be called the **level** of the group G.

3.4.1.2 Isogeny estimates

Our strategy for obtaining explicit estimates is a variant of the approach of [71] – cf. especially lemmas 3.1, 3.2 of *op. cit.* To ease the notation, when λ is a place of E we identify $\mathcal{O}_\lambda/\lambda$ with $\mathbb{F}_q$ for a suitable $q = \ell^f$. Also recall that we have introduced the residual representation G_λ in definition 3.4.4.

Lemma 3.4.12. *Suppose G_λ fixes a subspace Γ of dimension 1 of $\mathbb{F}_q^2$. Then $\ell \leq b_0(A/K)$.*

Proof. Γ is fixed by G_λ and therefore defined over K. Consider the K-variety $A^* = A/\Gamma$, which comes equipped with a natural isogeny $\pi : A \twoheadrightarrow A^*$ of degree $|\Gamma| = |\mathbb{F}_q| = \ell^f$. Choose a K-isogeny $\psi : A^* \to A$ of degree $b \leq b_0(A/K)$. The composition $\psi \circ \pi$ is an endomorphism of A, so by hypothesis it is given by a certain $e \in \mathrm{End}(A) \subseteq \mathcal{O}_E$. Now, as e kills Γ and $\mathrm{Ann}(\Gamma) = \lambda$, we must have $\lambda \mid e$ (that is, e belongs to λ, where we identify a place with its corresponding prime ideal). It follows that $d := \deg(e) = N_{E/\mathbb{Q}}(e)^2$ (for this equality cf. [11], Chapter 5, Corollary 1.3) is divisible by $N_{E/\mathbb{Q}}(\lambda)^2$, which is just $|\mathbb{F}_q|^2 = \ell^{2f}$. Comparing degrees we have $\ell^{2f} \mid d = b\ell^f$, so ℓ^f divides b which, in turn, is at most $b_0(A/K)$. □

Similarly, an easy variant of the argument of lemma 3.3.17 gives

Lemma 3.4.13. *Suppose G_λ is commutative. Then $\ell^2 \leq b_0\left(A^2/K\right)$.*

3.4.1.3 Explicit bounds: split primes

In this section we consider those primes ℓ that split completely in E. The group H_ℓ of the following definition is the natural candidate for the image of ρ_ℓ, for $\ell \gg 0$: it is the largest (connected) group whose elements are simultaneously symplectic isometries for the Weil pairing and contained in the centralizer of the action of E.

Definition 3.4.14. Let ℓ be a prime that splits completely in E. We set

$$H_\ell = \left\{ (h_\lambda)_{\lambda|\ell} \in \prod_{\lambda|\ell} \mathrm{GL}_2(\mathcal{O}/\lambda) \;\middle|\; \det(h_{\lambda_1}) = \det(h_{\lambda_2}) \in \mathbb{F}_\ell^\times \ \forall \lambda_1, \lambda_2 | \ell \right\},$$

where the product is over the places of E that divide ℓ.

Lemma 3.4.15. *For any split prime ℓ, the group G_ℓ is contained in H_ℓ.*

Proof. The determinant of every ρ_λ agrees with the cyclotomic character (lemma 3.4.5), so any two h_λ's will have the same determinant. □

With this notation, the bound we obtain is as follows:

Theorem 3.4.16. *Let A/K be an abelian variety whose endomorphism algebra $\mathrm{End}_{\overline{K}}(A) \otimes \mathbb{Q}$ is a totally real number field E of degree equal to $\dim A$. Suppose that the action of E is defined over K. If ℓ is a prime that does not ramify in K, is completely split in E and is strictly larger than $b(2[K:\mathbb{Q}], 2\dim(A), 2h(A))^{1/2}$, then the equality $G_\ell = H_\ell$ holds.*

To make the notation lighter we introduce the following definition:

Definition 3.4.17. Let A/K be an abelian variety. We set

$$M(A/K) := b(2[K:\mathbb{Q}], 2\dim(A), 2h(A))^{1/2}.$$

Lemma 3.4.18. *If ℓ is a rational prime larger than $M(A/K)$ and λ is a place of E above ℓ, then the group G_λ contains $\mathrm{SL}(2, \mathbb{F}_\ell)$.*

Proof. Let ℓ be a prime for which G_λ does not contain $\mathrm{SL}(2,\mathbb{F}_\ell)$ and recall that we identify $\mathcal{O}_\lambda/\lambda$ with $\mathbb{F}_q$ for a suitable $q=\ell^f$. By the Dickson classification (theorem 3.4.10; cf. also [116, §2]) we know that if G_λ does not contain $\mathrm{SL}(2,\mathbb{F}_\ell)$, then the following are the only possibilities:

(I) G_λ is contained in a Borel subgroup of $\mathrm{GL}(2,\mathbb{F}_q)$: by definition, such a subgroup fixes a line, therefore $\ell \leq b(A/K)$ by lemma 3.4.12.

(II) G_λ is contained in a Cartan subgroup of $\mathrm{GL}(2,\mathbb{F}_q)$: then $\ell^2 \leq b\left(A^2/K\right)$ by lemma 3.4.13.

(III) G_λ is contained in the normalizer of a Cartan subgroup of $\mathrm{GL}(2,\mathbb{F}_q)$: let C be this Cartan subgroup and N its normalizer. By the Dickson classification, the index $[N:C]$ is 2, so the morphism

$$\Gamma_K \to G_\lambda \to \frac{G_\lambda}{G_\lambda \cap C} \hookrightarrow \frac{N}{C}$$

induces a quadratic character of $\mathrm{Gal}(\overline{K}/K)$. The kernel of this character is associated with a field extension K'/K that satisfies $[K':K] \leq |N/C| = 2$. By construction, the image of $\mathrm{Gal}(\overline{K'}/K')$ in $\mathrm{Aut}\,(A[\lambda])$ is contained in C, so applying lemma 3.4.13 to $A_{K'}$ we see that ℓ is at most $b\left(A^2/K'\right)^{1/2} \leq b(2[K:\mathbb{Q}], 2\dim(A), 2h(A))^{1/2}$.

(IV) The projective image $\mathbb{P}G_\lambda$ of G_λ is a finite group of order at most 60: by lemma 3.4.12 we have $\ell \leq b(A/K'')$, where K'' is the field associated with the kernel of $\mathrm{Gal}(\overline{K}/K) \to G_\lambda \to \mathbb{P}G_\lambda$.

It is clear that G_λ does not fall in any of the previous cases – and therefore contains $\mathrm{SL}_2(\mathbb{F}_\ell)$ – as soon as ℓ is larger than $\max\left\{b(A/K), b\left(A^2/K\right)^{1/2}, b(A^2/K')^{1/2}, b(A/K'')\right\}$. It is immediate to check that this maximum is at most $M(A/K)$. □

Corollary 3.4.19. *Let ℓ be a rational prime that is unramified in K, completely split in E and strictly larger than $M(A/K)$. The group G_λ equals $\mathrm{GL}(2,\mathbb{F}_\ell)$ for every place λ of E dividing ℓ.*

The final piece we need to prove theorem 3.4.16 is the following lemma:

Lemma 3.4.20. *If $\ell > M(A/K)$ is totally split in E and does not ramify in K, and λ_1, λ_2 are two places of E dividing ℓ, then the projection $G_\ell \cap \mathrm{SL}(A[\ell]) \to \mathrm{SL}(\mathbb{F}_{\lambda_1}) \times \mathrm{SL}(\mathbb{F}_{\lambda_2})$ is surjective.*

Proof. Let $\ell > M(A/K)$ be a rational prime that is totally split in E and λ_1, λ_2 be places of E lying over ℓ. As $\mathrm{SL}_2(\mathbb{F}_\ell)$ does not have any nontrivial abelian quotients for $\ell \geq 5$, lemma 3.4.8 implies that

$$G_\ell \cap \mathrm{SL}(A[\ell]) = \prod_{\lambda|\ell} \mathrm{SL}_2(\mathbb{F}_\ell)$$

if and only if for every pair of different places λ_1, λ_2 of E above ℓ the projection of $G_\ell \cap \mathrm{SL}(A[\ell])$ to $\mathrm{SL}_2(\mathbb{F}_{\lambda_1}) \times \mathrm{SL}(\mathbb{F}_{\lambda_2})$ is surjective. Let

$$D = \left\{(b, b') \in \mathrm{GL}(F_{\lambda_1}) \times \mathrm{GL}(F_{\lambda_2}) \;\middle|\; \det(b) = \det(b')\right\},$$

and assume that $G_\ell \to D$ is not surjective (this is even more general than the statement we actually need). We want to derive a contradiction. Let f, χ be the morphisms given by lemma 3.4.9 when

applied to the image of G_ℓ in D (the hypotheses of this lemma are satisfied thanks to corollary 3.4.19). Assume first that $\chi \equiv 1$. Let Γ be the subgroup of $A[\ell]$ given by

$$\left\{ (x, y, 0, \dots, 0) \in A[\lambda_1] \times A[\lambda_2] \times \prod_{\substack{\lambda \mid \ell \\ \lambda \neq \lambda_1, \lambda_2}} A[\lambda] \;\middle|\; y = f(x) \right\}$$

and let $A^* = A/\Gamma$. Denote by π the canonical projection $A \to A^*$ and let φ be an isogeny $A^* \to A$, of degree b bounded by $b(A/K)$, as guaranteed by theorem 3.2.2. The composition $\varphi \circ \pi : A \to A$ is given by a certain $e \in \mathcal{O}_E$, and it kills Γ. In particular, e must be divisible by both λ_1 and λ_2. Indeed, e acts on $A[\lambda_1]$ via e_1, the class of e in $\mathbb{F}_{\lambda_1}$, and via e_2, the class of e in $\mathbb{F}_{\lambda_2}$, on $A[\lambda_2]$. If $x \in A[\lambda_1]$ is any element of order ℓ, then fx has the same property (since f is an isomorphism), so $(e_1, e_2) \cdot (x, fx) = (e_1, e_2 f(x))$ vanishes if and only if both e_1 and e_2 do. We deduce

$$\ell^4 = N_{K/\mathbb{Q}}(\lambda_1 \lambda_2)^2 \mid N_{K/\mathbb{Q}}(e)^2 = \deg(e) = b \cdot |\Gamma| = b\ell^2,$$

so $\ell \leq b^{1/2} \leq b(A/K)^{1/2}$. On the other hand, if χ is not identically 1, then its kernel defines a quadratic extension K' of K for which $\chi \circ \rho_{\lambda_1} \equiv 1$, therefore applying the same argument to K' we deduce $\ell \leq b(2[K : \mathbb{Q}], \dim(A), h(A))^{1/2}$. It is immediate to check that this number is smaller than $M(A/K)$, and the lemma follows. □

The main result of this section is now well within our reach:

Proof. (of theorem 3.4.16) Let ℓ be completely split in E, unramified in K and larger than $M(A/K)$. By the previous lemma we have $G_\ell \cap SL(A[\ell]) = \prod_{\lambda \mid \ell} \mathrm{SL}_2(\mathbb{F}_\ell)$, and by surjectivity of the determinant (guaranteed by lemma 3.4.6) this means $G_\ell = H_\ell$. □

3.4.1.4 Explicit bounds: non-split primes

Let ℓ be a prime unramified in E, and write $\prod_{i=1}^n \lambda_i$ for its factorization in $\mathcal{O}_E$. Our next aim is to show that, for every λ_i lying above ℓ, the group G_{λ_i} contains $\mathrm{SL}_2(\mathbb{F}_{\lambda_i})$.
Assume that $\ell > M(A/K) > 5$, so that by lemma 3.4.18 we know that every G_{λ_i} contains $\mathrm{SL}_2(\mathbb{F}_\ell)$. We set $\beta_i = [\mathbb{F}_{\lambda_i} : \mathbb{F}_\ell]$, and notice that ℓ^{β_i} is the order of the residue field at λ. The assumption that G_{λ_i} contains $\mathrm{SL}_2(\mathbb{F}_\ell)$ immediately implies that G_{λ_i} must be of type (5a) or (5b) in the notation of theorem 3.4.10. Suppose first G_{λ_i} is of type (5a), generated (up to conjugation) by $\mathrm{SL}_2(\mathbb{F}_{\ell^{\alpha_i}})$ and by a scalar matrix $V = \mu \cdot \mathrm{Id}$. Since the determinant of any element in G_{λ_i} lies in $\mathbb{F}_\ell^\times$ we know that $\det V = \mu^2$ is an element of $\mathbb{F}_\ell$, hence $V^2 \in \mathrm{GL}_2(\mathbb{F}_\ell)$. In particular, G_{λ_i} contains as a subgroup of index 2 the group generated by $\mathrm{SL}_2(\mathbb{F}_{\ell^{\alpha_i}})$ and V^2, which is a subgroup of $\mathrm{GL}_2(\mathbb{F}_{\ell^{\alpha_i}})$. Furthermore, if G_{λ_i} is of type (5b), then it contains a group of type (5a) as a subgroup of index 2. We thus deduce:

Lemma 3.4.21. *Let $\ell > M(A/K)$, so that G_{λ_i} is of type (5a) or (5b). Let α_i be the level of G_{λ_i}. There exists an extension K' of K, of degree at most 4, such that – up to conjugation – the image of $\rho_{\lambda_i} : \mathrm{Gal}(\overline{K'}/K') \to \mathrm{GL}_2(\mathbb{F}_{\lambda_i})$ is contained in $\mathrm{GL}_2(\mathbb{F}_{\ell^{\alpha_i}})$.*

Our aim is to show that – at least for ℓ large enough – the level α_i must necessarily equal β_i, the degree $[\mathbb{F}_{\lambda_i} : \mathbb{F}_\ell]$:

Lemma 3.4.22. *Suppose that, up to conjugation in $\mathrm{GL}_2(\mathbb{F}_{\lambda_i})$, the group G_{λ_i} contains $\mathrm{SL}_2(\mathbb{F}_\ell)$ and is contained in $\mathrm{GL}_2(\mathbb{F}_{\ell^{\alpha_i}})$ for some $\alpha_i < \beta_i$. Then $\ell \leq b_0(A/K)^{1/2}$.*

Proof. For every place λ of E above ℓ we can identify $A[\lambda]$ with $\mathbb{F}_\lambda^2$, and for the factor $A[\lambda_i]$ the hypothesis allows us to choose coordinates in such a way that the image of ρ_{λ_i} is contained $\mathrm{GL}_2(\mathbb{F}_{\ell^{\alpha_i}})$. Consider now the subspace of $A[\ell]$ given by

$$\Gamma = \left\{ (x_\lambda) \in \prod_{\lambda|\ell} A[\lambda] \cong \prod_{\lambda|\ell} \mathbb{F}_\lambda^2 \;\middle|\; x_{\lambda_i} \in (\mathbb{F}_{\ell^{\alpha_i}})^2, x_\lambda = 0 \text{ for } \lambda \neq \lambda_i \right\}.$$

Notice that, by construction, Γ contains torsion points whose annihilator in $\mathrm{End}(A)$ is exactly (λ_i), for example the point whose coordinates (in our basis) are $(1,0)$ along the λ_i-component. Furthermore, the subgroup Γ is Galois-stable: indeed, for any $g \in \mathrm{Gal}(\overline{K}/K)$ and every $(x_\lambda) \in \Gamma$, we have

$$(\rho_\lambda(g) \cdot x_\lambda)_{\lambda_i} = \rho_{\lambda_i}(g) \cdot x_{\lambda_i} \in \mathbb{F}_{\ell^{\alpha_i}}^2,$$

since both the coefficients of the vector x_{λ_i} and those of the matrix $\rho_{\lambda_i}(g)$ lie in $\mathbb{F}_{\ell^{\alpha_i}}$. It follows that the abelian variety $A' = A/\Gamma$ is defined over K, and there are isogenies $\pi : A \to A'$ (the canonical projection, of degree $\ell^{2\alpha_i}$) and $\psi : A' \to A$ (which can be chosen to be of degree at most $b_0(A/K)$). Notice now that $\psi \circ \pi$ is an endomorphism $e \in \mathcal{O}_E$ of A, and it kills a point whose annihilator is (λ). It follows that $(\lambda) \mid e$, so the degree of $\psi \circ \pi$ satisfies

$$\ell^{2\beta_i} = N_{E/\mathbb{Q}}(\lambda)^2 \mid N_{E/\mathbb{Q}}(e)^2 \mid \deg(\psi \circ \pi) = \ell^{2\alpha_i} \deg \psi \leq \ell^{2\alpha_i} b_0(A/K),$$

hence $\ell^2 \leq \ell^{2(\beta_i - \alpha_i)} \leq b_0(A/K)$. The lemma follows. □

Combining the previous two lemmas we find

Corollary 3.4.23. *Let $\ell > M(A/K)$ be a prime number and λ be a place of E above ℓ. The image of the representation $\rho_\lambda : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_2(\mathbb{F}_\lambda)$ contains $\mathrm{SL}_2(\mathbb{F}_\lambda)$.*

Proof. By lemma 3.4.21 we know that G_λ is of type (5a) or (5b) in the sense of theorem 3.4.10. Let α_i be the level of G_λ; it is clear that it is enough to show $\alpha_i = \beta_i$. By lemma 3.4.21, passing to an extension K' of K of degree at most 4 we can assume that (up to conjugation) the image of $\rho_{\lambda_i} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\, A[\lambda]$ is contained in $\mathrm{GL}_2(\mathbb{F}_{\ell^\alpha})$. The corollary then follows from lemma 3.4.22 (applied to A/K') and the obvious inequality $M(A/K) > b(A/K')^{1/2} \geq b_0(A/K')^{1/2}$. □

Lemma 3.4.24. *Let λ_1, λ_2 be two places of $\mathcal{O}_E$ above the prime $\ell \geq 5$. Suppose that $\ell > M(A/K)$: then the image of $\mathrm{Gal}(\overline{K}/K) \xrightarrow{\rho_{\lambda_1} \times \rho_{\lambda_2}} G_{\lambda_1} \times G_{\lambda_2}$ contains $\mathrm{SL}_2(\mathbb{F}_{\lambda_1}) \times \mathrm{SL}_2(\mathbb{F}_{\lambda_2})$.*

Proof. By corollary 3.4.23 G_{λ_i} contains $\mathrm{SL}_2(\mathbb{F}_{\lambda_i})$ for $i = 1, 2$. Let S be the image of G_ℓ in $G_{\lambda_1} \times G_{\lambda_2}$, and for the sake of simplicity write $S_i = \mathrm{SL}_2(\mathbb{F}_{\lambda_i})$ for $i = 1, 2$ and set $S^1 = S \cap (S_1 \times S_2)$. The claim of the lemma amounts to saying that $S^1 = S_1 \times S_2$. Suppose that this is not the case: then by Goursat's lemma there exist normal subgroups N_1, N_2 (of S_1, S_2 respectively) and an isomorphism $\varphi : S_1/N_1 \to S_2/N_2$ such that S^1 projects to the graph of φ in $S_1/N_1 \times S_2/N_2$. Comparing the orders of S_1/N_1 and S_2/N_2 (or, more precisely, their valuations at ℓ) easily gives $\mathbb{F}_{\lambda_1} = \mathbb{F}_{\lambda_2}$. We

can then deduce from lemma 3.4.9 the existence of an isomorphism $f : \mathbb{F}_{\lambda_1}^2 \to \mathbb{F}_{\lambda_2}^2$, of a character $\chi : S \to \{\pm 1\}$, and of an automorphism σ of $\mathbb{F}_{\lambda_1} = \mathbb{F}_{\lambda_2}$ such that $g_2 = \chi((g_1, g_2))\sigma\left(f g_1 f^{-1}\right)$ for all (g_1, g_2) in S. Assume first that $\chi \equiv 1$: then the subspace

$$\Gamma := \left\{ (x_\lambda) \in \prod_{\lambda|\ell} \mathbb{F}_\lambda^2 \cong \prod_{\lambda|\ell} A[\lambda] \;\middle|\; x_{\lambda_2} = \sigma(f x_{\lambda_1}),\ x_\lambda = 0 \text{ for } \lambda \neq \lambda_1, \lambda_2 \right\}$$

is Galois invariant, so the abelian variety $A^* := A/\Gamma$ is defined over K. Let $\pi : A \to A^*$ be the canonical projection and $\psi : A^* \to A$ be an isogeny of degree at most $b_0(A/K)$. Since Γ contains points whose annihilator is $(\lambda_1\lambda_2)$, it follows that $\psi \circ \pi =: e \in \mathcal{O}_E$ must be divisible by both λ_1 and λ_2. Hence if β denotes the common degree $[\mathbb{F}_{\lambda_1} : \mathbb{F}_\ell] = [\mathbb{F}_{\lambda_2} : \mathbb{F}_\ell]$ we have

$$\ell^{4\beta} = N_{E/\mathbb{Q}}\,(\lambda_1\lambda_2)^2 \leq \deg e = \deg\,(\psi \circ \pi) = \ell^{2\beta} \deg \psi \leq \ell^{2\beta} b_0(A/K),$$

whence $\ell \leq b_0(A/K)^{1/2} < M(A/K)$. If, on the other hand, χ is not the trivial character, then the kernel of $\mathrm{Gal}(\overline{K}/K) \to G_\ell \to S \xrightarrow{\chi} \{\pm 1\}$ defines an extension K' of K of degree 2, and repeating the same argument over K' we find $\ell \leq b_0(A/K; 2)^{1/2} < M(A/K)$. □

We are now ready to prove theorem 3.1.4, whose statement we reproduce here for the reader's convenience:

Theorem 3.4.25. *(Theorem 3.1.4) Let A/K be an abelian variety of dimension g. Suppose that $R = \mathrm{End}_{\overline{K}}(A)$ is an order in a totally real field E of degree g over $\mathbb{Q}$ (that is to say, A is of GL_2-type) and that all endomorphisms of A are defined over K. Let ℓ be a prime unramified both in K and in E and strictly larger than both $M(A/K)$ and $b(A/K)^g$: we have*

$$G_{\ell^\infty} = \left\{ x \in \mathrm{GL}_2\,(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}.$$

Proof. By lemma 3.4.24, the inequality imposed on ℓ guarantees that for every pair of places λ_1, λ_2 of E lying above ℓ the image of

$$\mathrm{Gal}(\overline{K}/K) \xrightarrow{\rho_{\lambda_1} \times \rho_{\lambda_2}} \mathrm{GL}_2(\mathbb{F}_{\lambda_1}) \times \mathrm{GL}_2(\mathbb{F}_{\lambda_2})$$

contains $\mathrm{SL}_2(\mathbb{F}_{\lambda_1}) \times \mathrm{SL}_2(\mathbb{F}_{\lambda_2})$. Since a group of the form $\mathrm{SL}_2(\mathbb{F}_\lambda)$ has no nontrivial abelian quotients (we can clearly assume $\ell \geq 5$), lemma 3.4.8 guarantees that G_ℓ contains $\prod_{\lambda|\ell} \mathrm{SL}_2(\mathbb{F}_\lambda)$, and proposition 3.4.7 then implies that G_{ℓ^∞} contains $\mathrm{SL}_2(\mathcal{O}_E \otimes \mathbb{Z}_\ell)$. Since furthermore the map $\det_{\mathcal{O}_E \otimes \mathbb{Z}_\ell} : G_{\ell^\infty} \to \mathbb{Z}_\ell^\times$ is surjective by lemma 3.4.6 (notice that ℓ is a good prime by proposition 3.4.2) we conclude that G_{ℓ^∞} contains $\left\{ x \in \mathrm{GL}_2\,(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}$, hence it is equal to it. □

Remark 3.4.26. It is not hard to show that when g is large enough we have $M(A/K) < b(A/K)^g$; in fact, $g \geq 33$ suffices.

The case of abelian surfaces follows at once:

Corollary 3.4.27. *(Corollary 3.1.5) Let A/K be an abelian surface. Suppose that $R = \mathrm{End}_{\overline{K}}(A)$ is an order in a real quadratic field E and that all endomorphisms of A are defined over K. Let ℓ be a rational prime, unramified both in K and in E and strictly larger than $b(2[K : \mathbb{Q}], 4, 2h(A))^{1/2}$: then we have*

$$G_{\ell^\infty} = \left\{ x \in \mathrm{GL}_2\,(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \;\middle|\; \det_{\mathcal{O}_E} x \in \mathbb{Z}_\ell^\times \right\}.$$

Proof. Immediate from the previous theorem and the (easy) inequality $M(A/K) > b(A/K)^2$. □

3.5 Type II – Quaternionic multiplication

In this section we establish the surjectivity result when the endomorphism ring of A, $\mathrm{End}_{\overline{K}}(A)$, is an order R in an indefinite (division) quaternion algebra D over $\mathbb{Q}$, and the action of R is defined over K. We let Δ be the discriminant of R.
We start by recalling a result from [6], cf. in particular Theorem 5.4 and the remarks preceding it.

Theorem 3.5.1. *Let ℓ be a prime not dividing Δ. Suppose that ℓ does not divide the degree of a fixed K-polarization of A. There exists a $\mathrm{Gal}(\overline{K}/K)$-equivariant isomorphism*

$$T_\ell(A) \cong W_{\ell^\infty} \oplus W_{\ell^\infty},$$

where W_{ℓ^∞} is a simple $\mathrm{Gal}(\overline{K}/K)$-module, free of rank 2 over $\mathbb{Z}_\ell$, equipped with a nondegenerate, $\mathrm{Gal}(\overline{K}/K)$-equivariant bilinear form

$$\langle \cdot, \cdot \rangle_{QM} : W_{\ell^\infty} \times W_{\ell^\infty} \to \mathbb{Z}_\ell(1).$$

Notation. We write W_ℓ for $W_{\ell^\infty}/\ell W_{\ell^\infty}$. It is a $\mathrm{Gal}(\overline{K}/K)$-module, free of rank 2 over $\mathbb{F}_\ell$.

Choosing bases for W_{ℓ^∞} and W_ℓ we have:

Lemma 3.5.2. *If ℓ does not divide Δ then G_ℓ can be identified with a subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$ (acting on $M_2(\mathbb{F}_\ell)$ on the right), and similarly G_{ℓ^∞} can be identified with a subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (acting on $M_2(\mathbb{Z}_\ell)$ on the right).*

In the light of the above lemma, we can consider G_ℓ as being a subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$, acting on $\mathbb{F}_\ell^4$ as two copies of the standard representation.

Lemma 3.5.3. *Suppose ℓ does not divide Δ and is larger than $b(2[K:\mathbb{Q}], 4, 2h(A))^{1/2}$. The group G_ℓ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$ under the above identification.*

Proof. This is a very minor variant of lemmas 3.4.12 and 3.4.13, so we only sketch the proof. If G_ℓ does not contain $\mathrm{SL}_2(\mathbb{F}_\ell)$, then Dickson's classification (theorem 3.4.10) implies that one of the following holds:

- G_ℓ is contained in a Borel subgroup: we can find a line $\Gamma \subseteq W_\ell$ that is stable under the action of G_ℓ. Applying an obvious variant of the argument of lemma 3.4.12 to the isogeny $A \to \frac{A}{\Gamma \oplus \Gamma}$ we find $\ell^2 \leq b(A/K)$.

- The projective image of G_ℓ has cardinality at most 60: by replacing K with an extension of degree at most 60 we are back to the previous case, and therefore $\ell^2 \leq b(60[K:\mathbb{Q}], 2, h(A))$.

- Up to replacing K with an extension K' of degree at most 2, G_ℓ is commutative, but does not entirely consist of scalars (this case being covered by the first one). We can choose an $\alpha \in G_\ell$ which is not a scalar, and apply a variant the argument of lemma 3.4.13 to the isogeny given by the natural projection from $A \times A$ to its quotient by the subgroup

  $$\left\{ (x_1, y_1, x_2, y_2) \in W_\ell \oplus W_\ell \oplus W_\ell \oplus W_\ell \cong A[\ell] \times A[\ell] \;\middle|\; x_2 = \alpha x_1 \right\}.$$

  The conclusion is now $\ell^2 \leq b(A^2/K') \leq b(2[K:\mathbb{Q}], 4, 2h(A))$.

Comparing the various bounds thus obtained we see that $b(2[K:\mathbb{Q}],4,2h(A))^{1/2}$ is much larger than any of the others, thus establishing the lemma. □

Lemma 3.5.4. *Suppose ℓ is a prime that does not divide Δ, so that $R\otimes\mathbb{Z}_\ell \cong M_2(\mathbb{Z}_\ell)$. Suppose furthermore that ℓ does not divide the degree of a fixed K-polarization of A. For any $g\in \mathrm{Gal}(\overline{K}/K)$ the determinant of $\rho_\ell(g)$, thought of as an element of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (and not of $\mathrm{GSp}(T_\ell(A))$), is $\chi_\ell(g)$.*

Proof. This is the same argument as for elliptic curves. If we fix a basis e_1,e_2 of W_{ℓ^∞} and write $\begin{pmatrix} a & b \\ c & d\end{pmatrix}$ for the matrix representing the action of $\rho_{\ell^\infty}(g)$ in this basis, we obtain

$$\begin{aligned}\chi_\ell(g)\langle e_1,e_2\rangle_{QM} &= \langle \rho_{\ell^\infty}(g)e_1, \rho_{\ell^\infty}(g)e_2\rangle_{QM}\\ &= \langle ae_1+ce_2, be_1+de_2\rangle_{QM}\\ &= ad\langle e_1,e_2\rangle + bc\langle e_2,e_1\rangle_{QM}\\ &= (ad-bc)\langle e_1,e_2\rangle_{QM},\end{aligned}$$

and since $\langle e_1,e_2\rangle_{QM}$ does not vanish we obtain $\chi_\ell(g)=(ad-bc)=\det\rho_\ell(g)$ as claimed. □

Theorem 3.5.5. *Suppose that every endomorphism of A is defined over K. Suppose furthermore that ℓ does not divide Δ, does not ramify in K, and is strictly larger than $b(2[K:\mathbb{Q}],4,2h(A))^{1/2}$. Then $G_{\ell^\infty}=(R\otimes\mathbb{Z}_\ell)^\times$.*

Proof. As $b(2[K:\mathbb{Q}],4,2h(A))^{1/2} > b(A/K)$, by [28, Théorème 1] we see that ℓ does not divide the degree of a minimal polarization of A, so by theorem 3.5.1 we have well-defined modules W_{ℓ^∞}, W_ℓ and a nondegenerate bilinear form $\langle\cdot,\cdot\rangle_{QM}$.
By lemma 3.5.3 the inequality imposed on ℓ guarantees that G_ℓ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$. It follows that G_{ℓ^∞} is a closed subgroup of $(R\otimes_{\mathbb{Z}}\mathbb{Z}_\ell)^\times \cong \mathrm{GL}_2(\mathbb{Z}_\ell)$ whose projection modulo ℓ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$. Since we certainly have $\ell\geq 5$, it follows from lemma 3.3.16 that G_{ℓ^∞} contains $\mathrm{SL}_2(\mathbb{Z}_\ell)$. On the other hand, the previous lemma and the condition that ℓ is unramified in K ensure that $\det: G_{\ell^\infty}\to\mathbb{Z}_\ell^\times$ is onto, so $G_{\ell^\infty}=\mathrm{GL}_2(\mathbb{Z}_\ell)$ as claimed. □

Let us make a few closing remarks on this case. It is a general philosophy that – at the level of Galois representations – a variety of dimension $2g$ with quaternionic multiplication by an algebra D (whose center is the number field L) should behave like a variety of dimension g and endomorphism algebra L. The proof we have just given shows that this philosophy is very much correct in the case of surfaces, and indeed from lemma 3.5.3 onward this is virtually the same proof as for elliptic curves (cf. for example [71]). Even more precisely, write the bound we obtained for a surface in the form $b(2[K:\mathbb{Q}],2\dim(A),2h(A))^{1/2}$; for an elliptic curve E/K without (potential) complex multiplication, the Galois representation is surjective onto $\mathrm{GL}_2(\mathbb{Z}_\ell)$ for every prime ℓ that does not ramify in K and is larger than $b(2[K:\mathbb{Q}],2\dim E,2h(E))^{1/2}$ (cf. [71]), which is formally the same expression. On the other hand, the actual numerical dependence of the present result on the height of A is much worse than the analogous one for elliptic curves, due to the strong dependence of the function $b([K:\mathbb{Q}],\dim A,h(A))$ on the parameter $\dim A$.

Remark 3.5.6. In the light of this discussion, the reader might suspect that the methods of chapter 1 might be generalized to give results on the index of the *adelic* representation attached to A. We do not attempt this here, for doing so would entail giving a classification of the integral Lie subalgebras of *any* $\mathbb{Z}_\ell$-form of $\mathfrak{sl}_2$: indeed, such algebras appear when we try to study the precise structure of G_{ℓ^∞} for those ℓ's that divide Δ. The task of classifying such algebras seems rather daunting, given that the easier problem of studying the $\mathbb{Q}_\ell$-forms of $\mathfrak{sl}_2$ is already highly nontrivial.

3.6 The index of the endomorphism ring

Let A/K be an absolutely simple abelian variety. Its endomorphism ring $R = \mathrm{End}_{\overline{K}}(A)$ is an order in a finite-dimensional division algebra D over $\mathbb{Q}$, and we are interested in giving a bound on the index of R in any maximal order $\mathcal{O}_D$ containing it. Note that when D is a field there is a unique maximal order, which is just the usual ring of integers, but when D is not commutative the index $[\mathcal{O}_D : R]$ might a priori depend on the choice of $\mathcal{O}_D$. The following proposition shows that this is not the case:

Proposition 3.6.1. *Let L be a number field, D a central simple algebra over L and R an order of D. Let $\mathcal{O}_D$ be a maximal order in D containing R. The index $[\mathcal{O}_D : R]$ does not depend on the choice of $\mathcal{O}_D$.*

Proof. Note first that any maximal order of D is stable under multiplication by $\mathcal{O}_L$ (indeed if S is a subring of D then the $\mathcal{O}_L$-module generated by S is again a subring of D), so the order R' generated by R and $\mathcal{O}_L$ is again contained in $\mathcal{O}_D$. We have $[\mathcal{O}_D : R] = [\mathcal{O}_D : R'][R' : R]$, and since $[R' : R]$ clearly does not depend on $\mathcal{O}_D$ we can assume that $R = R'$, i.e. that R is an $\mathcal{O}_L$-order. Under this additional assumption we have

$$\mathcal{O}_D/R \cong \bigoplus_{v \text{ finite place of } L} \frac{\mathcal{O}_D \otimes \mathcal{O}_{L_v}}{R \otimes \mathcal{O}_{L_v}},$$

so that $[\mathcal{O}_D : R] = \prod_{v \text{ finite place of } L}[\mathcal{O}_D \otimes \mathcal{O}_{L_v} : R_v]$, where $R_v = R \otimes \mathcal{O}_{L_v}$. It is then clear that $\mathcal{O}_D \otimes \mathcal{O}_{L_v}$ is a maximal order in $\mathcal{O}_D \otimes L_v$, and that it is enough to prove that at every finite place the index $[\mathcal{O}_D \otimes \mathcal{O}_{L_v} : R_v]$ is independent of the choice of $\mathcal{O}_D$. We are thus reduced to the local complete case, so Theorem 17.3 of [107] applies to give that all maximal orders in $\mathcal{O}_D \otimes L_v$ are conjugated. We now write the index $[\mathcal{O}_D \otimes \mathcal{O}_{L_v} : R_v]$ as the ratio $\dfrac{covol(R_v)}{covol(\mathcal{O}_D \otimes \mathcal{O}_{L_v})}$, where the covolume is taken with respect to any Haar measure (on $\mathcal{O}_D \otimes L_v$): as the Haar measure is invariant under conjugation, this quantity does not depend on $\mathcal{O}_D$. □

In order to simplify matters it is convenient to assume that all the endomorphisms of A are defined over K. This condition is completely harmless, since it can be achieved at the expenses of a controllable extension of K:

Lemma 3.6.2. *([129, Theorem 4.1]) There exists a number field K', with $[K' : K]$ bounded only in terms of $g = \dim(A)$, such that all the endomorphisms of A are defined over K'. We can take $[K' : K] \leq 2 \cdot (9g)^{2g}$.*

From now on we will therefore assume that all the endomorphisms of A are already defined over K. In order to get estimates in the case of noncommutative endomorphism algebras we will need the following lemma, which is essentially [143, proposition 2.5.4]: even though the latter was stated only for commutative endormorphism rings, the same proof works in the general case as well.

Lemma 3.6.3. *Let D be a division algebra, $R \subseteq S$ be orders in D and A/K be an abelian variety with $\mathrm{End}_K(A) = R$. There exists an abelian variety B/K, isogenous to A over K, such that $\mathrm{End}_K(B) \supseteq S$.*

Corollary 3.6.4. *Let A/K be an Abelian variety with endomorphism ring R, $D = R \otimes \mathbb{Q}$, and $\mathcal{O}_D$ any maximal order containing R. Suppose that all the endomorphisms of A are defined over K. There exists an Abelian variety A'/K and two isogenies $\varepsilon_1 : A \to A', \varepsilon_2 : A' \to A$, defined over K, such that $\mathrm{End}(A') = \mathcal{O}_D$ and*

$$\max\{\deg(\varepsilon_1), \deg(\varepsilon_2)\} \leq b(A/K).$$

Proof. Lemma 3.6.3 shows the existence of a K-variety A' having $\mathcal{O}_D$ as its endomorphism ring, so the claim follows from [28, Theorem 1.4] (which is a symmetric version of theorem 3.2.2, bounding degrees of minimal isogenies both from A to A' and from A' to A). □

We can now deduce the desired bound on $[\mathcal{O}_D : R]$:

Proposition 3.6.5. *The inequality $[\mathcal{O}_D : R] \leq b(A/K)^{\dim_{\mathbb{Q}}(D)}$ holds.*

Proof. Let $A', \varepsilon_1, \varepsilon_2$ be as in the above corollary. Consider the following linear map:

$$\begin{array}{rccccc} \varphi: & \mathrm{End}(A') & \to & \mathrm{End}(A) & \hookrightarrow & \mathrm{End}(A') \\ & e & \mapsto & \varepsilon_2 \circ e \circ \varepsilon_1, & & \end{array}$$

where the second embedding is given by the fact that $R = \mathrm{End}(A)$ is an order in D and $\mathcal{O}_D$ is a maximal order containing R. Note that $\mathrm{End}(A)$ is endowed with a positive-definite quadratic form given by the degree. We consider $\mathrm{End}(A)$ and $\mathrm{End}(A')$ both as lattices inside $D_{\mathbb{R}} = \mathrm{End}(A') \otimes_{\mathbb{Z}} \mathbb{R}$, and observe that the degree map extends naturally to a positive-definite quadratic form on $D_{\mathbb{R}}$. This makes $D_{\mathbb{R}}$ into an Euclidean space, which in particular comes equipped with a natural (Lebesgue, say) measure. Denote r the dimension of $D_{\mathbb{R}}$, which is also the dimension of D as a $\mathbb{Q}$-vector space. As $\deg(e_1 \circ e_2) = \deg(e_1) \cdot \deg(e_2)$ for any pair of isogenies between abelian varieties, we have

$$\deg(\varphi(e)) = \deg(\varepsilon_2 \circ e \circ \varepsilon_1) = \deg(\varepsilon_1)\deg(\varepsilon_2)\deg(e) \leq b(A/K)^2 \deg(e).$$

Extend φ by linearity to an endomorphism (which we still denote by φ) of $D_{\mathbb{R}}$ and fix a deg-orthonormal basis $\gamma_1, \ldots, \gamma_r$ of $D_{\mathbb{R}}$. By construction $\varphi(\mathcal{O}_D) \subseteq R$, so that we have the inequality

$$[\mathcal{O}_D : R] = \frac{covol(R)}{covol(\mathcal{O}_D)} \leq \frac{covol\left(\varphi(\mathcal{O}_D)\right)}{covol(\mathcal{O}_D)} = \frac{\det(\varphi) covol(\mathcal{O}_D)}{covol(\mathcal{O}_D)} = \det(\varphi).$$

Write $\varphi(\gamma_i) = \sum_{j=1}^{r} a_{ij}\gamma_j$ with $a_{ij} \in \mathbb{R}$ for the matrix representing φ in the basis of the γ_j's. Let $\lambda(\cdot,\cdot)$ be the bilinear form associated with deg. Using the inequality $\deg(\varphi(e)) \leq b(A/K)^2 \deg(e)$ we deduce

$$\deg\left(\sum_j a_{ij}\gamma_j\right) = \deg(\varphi(\gamma_i)) \leq b(A/K)^2 \deg(\gamma_i) = b(A/K)^2 \quad \forall i = 1, \ldots, r,$$

so

$$b(A/K)^2 \geq \lambda\left(\sum_j a_{ij}\gamma_j, \sum_k a_{ik}\gamma_k\right) = \sum_{j,k} a_{ij}a_{ik}\lambda(\gamma_j,\gamma_k) = \sum_j a_{ij}^2 \quad \forall i = 1,\ldots,r :$$

equivalently, the L^2-norm of each row of the matrix (a_{ij}) is bounded by $b(A/K)$. Hadamard's inequality then gives

$$[\mathcal{O}_D : R] \leq \det(\varphi) \leq \prod_{i=1}^{r} ||a_i||_{L^2} \leq b(A/K)^r,$$

which is the desired estimate. □

Chapter 4

Abelian threefolds, and a glimpse into the higher-dimensional situation

4.1 Introduction

Let K be a number field and A be a K-abelian variety. The aim of the present chapter is to study the Galois representations attached to A, under the assumption that $\mathrm{End}_{\overline{K}}(A)$ is $\mathbb{Z}$ and $g = \dim A$ is an odd number greater than or equal to 3. More precisely, we are interested in the family of representations

$$\rho_{\ell^\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(T_\ell(A)) \cong \mathrm{GL}_{2g}(\mathbb{Z}_\ell)$$

arising (after a choice of basis) from the ℓ-adic Tate modules of A. We shall also consider the residual mod-ℓ representations

$$\rho_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(A[\ell]) \cong \mathrm{GL}_{2g}(\mathbb{F}_\ell),$$

and write G_{ℓ^∞} (resp. G_ℓ) for the image of ρ_{ℓ^∞} (resp. of ρ_ℓ). Under our assumptions, it is known by work of Serre [118] that for all ℓ large enough (with respect to A/K) the equality $G_{\ell^\infty} = \mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ holds. Our aim is to explicitly find a bound ℓ_0 (depending on A and K) such that, for all primes $\ell > \ell_0$, the representation ρ_{ℓ^∞} is onto $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$.
For technical reasons we need to impose an additional constraint on the dimensions g we take into account. We say that the odd number g satisfies condition $(*)$ if the following holds (cf. definition 4.3.13 for the notion of class-$\mathcal{S}$ subgroups):

> let $\ell > \frac{1}{2}(2g+1)^{12g}$ be a prime number, and let G be a class-$\mathcal{S}$ maximal subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ such that $\mathrm{soc}(\mathbb{P}G)$ is a simple group of Lie type: then $\mathrm{soc}(\mathbb{P}G) \cong \mathrm{PSL}_2(\mathbb{F}_\ell)$.

Remark 4.1.1. Condition $(*)$, albeit very unnatural, is at least not too severe a restriction: indeed we can show that the set $\mathcal{E}$ of odd numbers g that fail to satisfy it has density zero (theorem 4.10.1). Furthermore, as it will be clear from sections 4.10 and 4.11, there is an algorithmic procedure that allows us to decide whether a certain g has property $(*)$ or not: we use this procedure to show that $(*)$ holds for all odd numbers in the interval $3 \leq g \leq 100$ with the exception of $7, 55$ and 63 (proposition 4.11.3). Finally, it is very likely that condition $(*)$ is not necessary for our results to hold, but we are for now unable to get rid of it.

To state our results more compactly we introduce the following functions:

Definition 4.1.2. Let K be a number field and A/K be an abelian variety of dimension g. We let $\alpha(g) = 2^{10}g^3$ and define

$$b(A/K) = b([K:\mathbb{Q}], g, h(A)) = \left((14g)^{64g^2}[K:\mathbb{Q}]\max\left(h(A), \log[K:\mathbb{Q}], 1\right)^2\right)^{\alpha(g)},$$

$$b(A/K; d) = b(d[K:\mathbb{Q}], g, h(A)),$$

where $h(A)$ is the stable Faltings height of A.

Our first result is the following explicit surjectivity theorem:

Theorem 4.1.3. *Let A/K be an abelian variety of dimension g and G_{ℓ^∞} be the image of the natural representation $\rho_{\ell^\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\, T_\ell A$. Suppose that:*

1. $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$;
2. *$g \geq 3$ is an odd number satisfying condition $(*)$;*
3. *there exists a place v of K, of good reduction for A and with residue field of order q_v, such that the characteristic polynomial of the Frobenius at v acting on $T_\ell A$ has Galois group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^g \rtimes S_g$.*

The equality $G_{\ell^\infty} = \mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ holds for every prime ℓ unramified in K, strictly larger than

$$\max\left\{(2q_v)^{2g\cdot g!},\ b(A/K; g!),\ b(A^2/K; g)^{1/2g}\right\},$$

and such that there is a place of K of residue characteristic ℓ at which A has semistable reduction. Furthermore, the term $b(A^2/K; g)^{1/2g}$ can be omitted from the maximum if $g \geq 19$.

In practice, it is usually very easy to find a place v as in the statement of theorem 4.1.3 (see for example the explicit calculation of section 4.12 and the remarks preceding lemma 4.7.6); however, in order to have a completely effective result we also need to show that the number q_v can be effectively bounded *a priori* in terms of simple arithmetic invariants of A/K. While unfortunately we cannot do this for arbitrary g, for simple abelian *threefolds* we prove:

Theorem 4.1.4. *(Theorem 4.9.17) Let A/K be an abelian variety of dimension 3 such that $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. Denote by $\mathcal{N}^0_{A/K}$ the naive conductor of A/K, that is, the product of the prime ideals of $\mathcal{O}_K$ at which A has bad reduction, and suppose that $A[7]$ is defined over K.*

- *Assume the Generalized Riemann Hypothesis: then the equality $G_{\ell^\infty} = \mathrm{GSp}_6(\mathbb{Z}_\ell)$ holds for every prime ℓ unramified in K and strictly larger than $(2q)^{48}$, where*

$$q = b(A^2/K; 3)^8\left(\log|\Delta_{K/\mathbb{Q}}| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right)^2.$$

- *Unconditionally, the same conclusion holds with*

$$q = \exp\left(cb(A^2/K; 3)^8\left(\log|\Delta_K| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right)^2\right),$$

where c is an absolute, effectively computable constant.

Remark 4.1.5. The condition that the 7-torsion points of A are defined over K is not very restrictive, for it can be met by simply replacing K by $K(A[7])$, cf. remark 4.9.18.

Remark 4.1.6. Unpublished work of Winckler [144] shows that c can be taken to be 27175010. Furthermore, if A/K is a semistable abelian variety, then $\log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)$ is bounded above by $[K:\mathbb{Q}]\,(c_0 h(A)+c_1)$ for certain constants c_0, c_1 depending only on $[K:\mathbb{Q}]$ and on $\dim A$: this result is stated and proved in [37] (see especially Theorem 6.5 of *op. cit.*) for abelian varieties over function fields, but the same proof works equally well also over number fields (for a detailed proof in the number field case see also [96, Theorem 1.1]).

To conclude this introduction let us describe the organization of this chapter. After two sections of preliminaries (§ 4.2 and 4.3) we study the various classes of maximal proper subgroups G of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$, showing that – at least for ℓ large enough – G_ℓ cannot be contained in any such G. This occupies sections 4.4, 4.5, 4.6, and 4.7, each of which deals with a different kind of maximal subgroup. Next in §4.8 we prove theorem 4.1.3, while section 4.9 contains a proof of theorem 4.1.4. In sections 4.10 and 4.11 we use representation theory (both in positive characteristic and over $\mathbb{C}$) to show that property $(*)$ is typical, in that it is true for a set of of density 1 which contains in particular all the odd numbers up to 100 with the only exception of $g = 7, 55, 63$. Finally, section 4.12 contains an example of an abelian threefold for which the previous theorems enable us to establish explicit surjectivity results.

We say a few more words on the techniques used in sections 4.4 through 4.7. Three classes of maximal subgroups (traditionally dubbed "imprimitive", "reducible", and "field extension" cases) are dealt with in section 4.4 as an almost immediate consequence of the isogeny theorem of Masser and Wüstholz [72] [70] (the completely explicit version we employ being due to Gaudron and Rémond [28]). Other maximal subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ are closely related to the image of the $2g$-dimensional projective symplectic representation of $\mathrm{PGL}_2(\mathbb{F}_\ell)$, and in section 4.5 we show that, for ℓ sufficiently large, G_ℓ cannot be contained in such a subgroup: this is obtained by comparing purely group-theoretical information with Raynaud's description of the structure of $A[\ell]$ as a module over the inertia group at a place of characteristic ℓ. The same results of Raynaud are also used in section 4.6 to eliminate the possibility of G_ℓ being a small "exceptional" (or "constant") group: we obtain a lower bound on $|\mathbb{P}G_\ell|$ that is linear in ℓ (and essentially uniform in A), which – combined with results of Larsen-Pink and Collins – shows that the exceptional case does not arise for ℓ larger than a certain explicit function of g. Finally, the hardest case is that of G_ℓ being contained in a "tensor product" subgroup. In §4.7 we show how, given a place v as in hypothesis (3) of theorem 4.1.3, one can produce a finite set of integers whose divisors include all the primes for which G_ℓ is of tensor product type; this is inspired by an argument of Serre [120], but his use of the characteristic polynomial of Fr_v is almost completely replaced by a direct study of the multiplicative relations satisfied by its roots. These relations also form the main object of interest in §4.9, where we exploit their simple form and the manageable structure of the subgroups of $\mathrm{GO}_3(\mathbb{F}_\ell)$ to show how, if $\dim A = 3$, a careful application of Chebotarev's theorem yields an effective bound on the residual characteristic of a place v with the desired properties.

4.2 Preliminaries

4.2.1 The isogeny theorem

The result that makes all the explicit estimates possible is the following theorem, due to Masser and Wüstholz [70] [72] and made explicit by Gaudron and Rémond [28]:

Theorem 4.2.1. *(Isogeny Theorem, [28, Theorem 1.4]) Let A/K be an abelian variety. For every abelian variety A^* defined over K that is K-isogenous to A, there exists a K-isogeny $A^* \to A$ whose degree is bounded by $b(A/K)$ (cf. definition 4.1.2).*

It is very likely that the function $b(A/K)$ of definition 4.1.2 is not the best possible one. Let us then introduce another function $b_0(A/K)$, which is by definition the optimal isogeny bound:

Definition 4.2.2. Let A/K be an abelian variety. We denote by $b_0(A/K)$ the smallest natural number such that, for every other abelian variety B/K that is K-isogenous to A, there exists a K-isogeny $B \to A$ of degree at most $b_0(A/K)$. We set $b_0(A/K;d) = \max_{[K':K]\leq d} b_0(A/K')$, where the maximum is taken over the finite extensions of K of degree at most d.

It is clear that the isogeny theorem implies that $b_0(A/K)$ and $b_0(A/K;d)$ are finite, and that $b_0(A/K;d) \leq b(d[K:\mathbb{Q}], \dim A, h(A)) =: b(A/K;d)$. Whenever possible, we will state our results in terms of b_0 instead of b; in some situations, however, in order to avoid cumbersome expressions involving maxima we simply give bounds in terms of the function b.

4.2.2 Weil pairing, Serre's lifting lemma

Let $A^\vee$ be the dual variety of A and let $\langle\cdot,\cdot\rangle$ denote the Weil pairing on $A \times A^\vee$. We also let $\mathbb{Z}_\ell(1)$ be the 1-dimensional Galois module the action on which is given by the cyclotomic character $\chi_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathbb{Z}_\ell^\times$. For any choice of a polarization $\varphi : A \to A^\vee$, the composition

$$T_\ell(A) \times T_\ell(A) \xrightarrow{\mathrm{id}\times\varphi} T_\ell(A) \times T_\ell(A^\vee) \xrightarrow{\langle\cdot,\cdot\rangle} \mathbb{Z}_\ell(1)$$

equips the Tate module $T_\ell(A)$ with a Galois-equivariant, skew-symmetric form which we still denote by $\langle\cdot,\cdot\rangle$ and call the Weil pairing on $T_\ell(A)$. By Galois-equivariance of $\langle\cdot,\cdot\rangle$, every element h of the group G_{ℓ^∞} preserves the form $\langle\cdot,\cdot\rangle$ up to multiplication by a scalar factor (called the *multiplier* of h), so G_{ℓ^∞} is in fact contained in $\mathrm{GSp}(T_\ell(A)\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell, \langle\cdot,\cdot\rangle)$, the group of symplectic similitudes of $T_\ell(A)\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ with respect to $\langle\cdot,\cdot\rangle$. Notice that the multiplier of h need not be an ℓ-adic unit, whence the need to tensor by $\mathbb{Q}_\ell$. Suppose however that ℓ does not divide the degree of the polarization φ: then $\mathrm{id}\times\varphi$ induces an isomorphism between $T_\ell(A) \times T_\ell(A)$ and $T_\ell(A) \times T_\ell(A^\vee)$, from which one easily deduces that the multiplier of every $h \in G_{\ell^\infty}$ is an ℓ-adic unit. It follows that (for these primes) G_{ℓ^∞} is a subgroup of $\mathrm{GSp}(T_\ell(A), \langle\cdot,\cdot\rangle)$, so, after a choice of basis, we can consider G_{ℓ^∞} as being a subgroup of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$.
Fix now (once and for all) a polarization φ of A of minimal degree. By [28, Théorème 1.1] we see that $\deg\varphi \leq b(A/K)$, so (since we only work with primes strictly larger than this quantity) we can assume that G_{ℓ^∞} is a subgroup of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$. Moreover, for such values of ℓ the Weil pairing is nondegenerate on $A[\ell]$, so for all primes $\ell > b(A/K)$ the group G_ℓ is a subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$.

Combining this remark with the following well-known lemma, originally due to Serre, will allow us to only consider the residual mod-ℓ representation ρ_ℓ instead of the full ℓ-adic system ρ_{ℓ^∞}:

Lemma 4.2.3. *Let g be a positive integer, $\ell \geq 5$ be a prime and G be a closed subgroup of $\mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$. Suppose that G surjects onto $\mathrm{Sp}_{2g}(\mathbb{F}_\ell)$ by reduction modulo ℓ: then $G = \mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$. Likewise, let H be a closed subgroup of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ whose reduction modulo ℓ contains $\mathrm{Sp}_{2g}(\mathbb{F}_\ell)$: then $H' = \mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$.*

Proof. The first statement is [120, Lemma 1 on p. 52], cf. also Theorem 1.3 in [136]. The second part follows by applying the first to $G = H'$ and noticing that the reduction modulo ℓ of H' contains the derived subgroup of $\mathrm{Sp}_{2g}(\mathbb{F}_\ell)$ which, for $\ell \geq 5$, is $\mathrm{Sp}_{2g}(\mathbb{F}_\ell)$ itself. □

Corollary 4.2.4. *Let $\ell > b(A/K)$: then G_ℓ is contained in $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. Suppose ℓ does not ramify in K: then $\mathrm{Gal}(\overline{K}/K) \xrightarrow{\chi_\ell} \mathbb{Z}_\ell^\times$ is surjective. In particular, if $\ell > b(A/K)$ does not ramify in K, the inclusion $\mathrm{Sp}(A[\ell]) \subseteq G_\ell$ implies $G_{\ell^\infty} = \mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$.*

We conclude this section of preliminaries by underlining once more our working assumption that ℓ does not divide the degree of a minimal polarization: this is a minor technical point, but it is necessary for all of our discussion to make sense.

Assumption 4.2.5. In all the statements to follow, we make the implicit hypothesis that the prime ℓ does not divide the degree of a minimal polarization of A. In particular, this allows us to identifty G_ℓ (resp. G_{ℓ^∞}) to a subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ (resp. $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$).

4.3 Maximal subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$

Thanks to corollary 4.2.4 we see that in order to prove theorem 4.1.3 it is enough to show that the equality $G_\ell = \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ holds all ℓ larger than a certain explicit bound. It is therefore not surprising that we may need a description of the maximal (proper) subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$: the core of our argument will consist in showing that – for ℓ large enough – G_ℓ cannot be contained in any proper subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$, and hence it has to coincide with all of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. The purpose of this section is to introduce some notation and state theorem 4.3.14, which gives precisely such a classification of the maximal subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. Our main references for this section are [19] and [49].

4.3.1 Group theoretical preliminaries

We now lay down some definitions and recall facts from finite group theory that will be needed in what follows.

Definition 4.3.1. Let G be a finite group. The **socle** of G, denoted $\mathrm{soc}(G)$, is the subgroup of G generated by the non-trivial minimal normal subgroups of G.

Definition 4.3.2. A finite group G is said to be **almost simple** if its socle is a non-abelian simple group. In this case, if we let $S = \mathrm{soc}(G)$, we have $S \leq G \leq \mathrm{Aut}(S)$, and S is a normal subgroup of G.

Lemma 4.3.3. *An almost simple group G does not possess non-trivial normal solvable subgroups.*

Proof. Suppose a nontrivial normal solvable subgroup exists. Then the collection of such subgroups is nonempty, and there is a minimal normal subgroup N_0 of G that is solvable (a subgroup of a solvable group is itself solvable). The definition of $\mathrm{soc}(G)$ implies $N_0 \subset \mathrm{soc}(G)$, and moreover N_0 is normal in $\mathrm{soc}(G)$ since it is normal in G. By simplicity of $\mathrm{soc}(G)$ this forces $N_0 = \mathrm{soc}(G)$; however, the latter is simple non-abelian, hence in particular not solvable, contradiction. □

Lemma 4.3.4. *An almost-simple group has a unique non-trivial minimal normal subgroup, which coincides with its socle.*

Proof. Let N be a non-trivial minimal normal subgroup. We have $N \triangleleft \mathrm{soc}(G)$, and as the latter is simple this forces $N = \mathrm{soc}(G)$. □

Definition 4.3.5. Let S be a finite group. The group $\mathrm{Inn}(S)$ of **inner automorphisms** of S is the image of the map

$$\begin{array}{ccc} S & \to & \mathrm{Aut}(S) \\ g & \mapsto & \left(\begin{array}{ccccc} \varphi_g : & S & \to & S \\ & s & \mapsto & gsg^{-1} \end{array} \right). \end{array}$$

The group $\mathrm{Inn}(S)$ is a normal subgroup of $\mathrm{Aut}(S)$. The quotient $\mathrm{Aut}(S)/\mathrm{Inn}(S)$ is called the **group of outer automorphisms** of G, and is denoted by $\mathrm{Out}(S)$.

Definition 4.3.6. A group is said to be **perfect** if it equals its commutator subgroup. If H is a finite group we denote by H^∞ the first perfect group contained in the derived series of H; equivalently,

$$H^\infty = \bigcap_{i \geq 0} H^{(i)},$$

where $H^{(0)} = H$ and $H^{(i+1)} = [H^{(i)}, H^{(i)}]$.

Lemma 4.3.7. *If G is almost simple we have $\mathrm{soc}(G) = G^\infty$; in particular, $\mathrm{soc}(G)$ is perfect.*

Proof. This follows immediately from the fact that the outer automorphism group of a simple group is solvable ([19, Theorem 1.3.2]). □

4.3.2 Definition of the classical groups

We now recall various standard constructions that are frequently used in the theory of finite matrix groups. Let F be a finite field of characteristic different from 2 and n be an odd integer. The **group of orthogonal transformations** of F^n is

$$\mathrm{GO}_n(F) = \left\{ x \in M_n(F) \;\middle|\; x^t x = \mathrm{Id} \right\}.$$

We also define the **special orthogonal group** $\mathrm{SO}_n(F) = \left\{ x \in \mathrm{GO}_n(F) \;\middle|\; \det x = 1 \right\}$ and the **group of orthogonal similarities**

$$\mathrm{CGO}_n(F) = \left\{ x \in M_n(F) \;\middle|\; \exists \lambda \in F^\times \text{ such that } x^t x = \lambda\, \mathrm{Id} \right\}.$$

Remark 4.3.8. These definitions also make sense for even n: in this case, however, there are two non-isomorphic groups preserving two non-equivalent quadratic forms on F^n. We shall not need to deal with this case.

We shall also need to consider the groups $\Omega_n(F)$:

Definition 4.3.9. ([19, p. 29]) Let $n \geq 3$ be odd: the group $\Omega_n(F)$ is the unique subgroup of $\mathrm{SO}_n(F)$ of index 2.

Remark 4.3.10. The group $\Omega_n(F)$ is usually introduced as the kernel of the so-called spinor norm $\mathrm{SO}_n(\mathbb{F}_\ell) \to \{\pm 1\}$; the precise definition of the spinor norm, however, is somewhat convoluted, while the simpler definition 4.3.9 is perfectly suitable for our purposes. Also notice that for any finite field F of odd characteristic the groups $\mathbb{P}\Omega_3(F)$ and $\mathrm{PSL}_2(F)$ are isomorphic, cf. [19, Proposition 1.10.1].

Let now n be any positive integer. The **standard symplectic form** on F^{2n} is

$$\begin{aligned} \langle \cdot, \cdot \rangle : \quad F^{2n} \times F^{2n} &\to F \\ (v, w) &\mapsto v^t J w, \end{aligned}$$

where $J := \mathrm{antidiag}(\underbrace{1, \ldots, 1}_{n}, \underbrace{-1, \ldots, -1}_{n})$. We can then introduce the **group of symplectic transformations**,

$$\mathrm{Sp}_{2n}(F) = \left\{ x \in M_{2n}(F) \;\middle|\; x^t J x = J \right\},$$

and the **group of symplectic similarities**

$$\mathrm{GSp}_{2n}(F) = \left\{ x \in M_{2n}(F) \;\middle|\; \exists \lambda \in F^\times \text{ such that } x^t J x = \lambda J \right\}.$$

Let V_1, V_2 be two vector spaces over F. The **Kronecker product** of $g_1 \in \mathrm{GL}(V_1)$ and $g_2 \in \mathrm{GL}(V_2)$ is the endomorphism $g_1 \otimes g_2$ of $V_1 \otimes_F V_2$ which acts as $(g_1 \otimes g_2)(v_1 \otimes v_2) = (g_1 v_1) \otimes (g_2 v_2)$ on decomposable elements, for all $v_1 \in V_1$ and $v_2 \in V_2$. If G and H are subgroups of $\mathrm{GL}_m(F)$, $\mathrm{GL}_n(F)$ respectively, we write $G \otimes H$ for the quotient of $G \times H$ by the equivalence relation

$$(a, b) \sim (c, d) \text{ if and only if there exists } \lambda \in F^\times \text{ such that } c = \lambda a,\ d = \lambda^{-1} b.$$

The group $G \otimes H$ is in a natural way a subgroup of $\mathrm{GL}_{mn}(F)$, the inclusion being given by identifying $(g, h) \in G \times H / \sim$ with $g \otimes h \in \mathrm{GL}_{mn}(F)$: the definition of $\sim$ ensures that this identification is well defined ([19, Proposition 1.9.8]).
Finally, whenever G is a subgroup of a certain linear group $\mathrm{GL}_n(F)$, we write $\mathbb{P}G$ for the image of G in the quotient $\mathbb{P}\mathrm{GL}_n(F) := \dfrac{\mathrm{GL}_n(F)}{F^\times \cdot \mathrm{Id}}$. We break this convention only for the groups $\mathbb{P}\mathrm{SL}_n(F)$ and $\mathbb{P}\mathrm{GL}_n(F)$, which in homage to the tradition will be denoted simply by $\mathrm{PSL}_n(F)$ and $\mathrm{PGL}_n(F)$.

4.3.3 Maximal subgroups of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$

We are now in a position to recall the classification of the maximal subgroups of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$. For simplicity of exposition, and since this is the only case we will need, we assume from now on that both n and ℓ are odd. Before stating the classification theorem we need to define some of the **Aschbacher classes**; we start with the notion of m-decomposition:

Definition 4.3.11. Let ℓ be an odd prime and $m \geq 2$ be an integer. An m-decomposition of $\mathbb{F}_\ell^{2n}$ is the data of m subspaces $V_1, \ldots, V_m$ of $\mathbb{F}_\ell^{2n}$, each of dimension $\frac{2n}{m}$, such that

- the restriction of the standard symplectic form of $\mathbb{F}_\ell^{2n}$ to V_i is either nondegenerate for every $i = 1, \ldots, m$, or trivial for every $i = 1, \ldots, m$;

- $\mathbb{F}_\ell^{2n} \cong \bigoplus_{i=1}^m V_i$.

We can now define the first four Aschbacher classes; as the precise definition of class $\mathcal{C}_3$ is somewhat complicated (cf. [19, Definition 2.2.5]), we shall limit ourselves to giving the property that will be crucial to us.

Definition 4.3.12. A subgroup G of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ is said to be:

1. **reducible**, or of **class** $\mathcal{C}_1$, if it stabilizes a linear subspace of $\mathbb{F}_\ell^{2n}$;
2. **imprimitive**, or of **class** $\mathcal{C}_2$, if there exists an m-decomposition $V_1, \dots, V_m$ which is stable under the action of G (i.e. for all $g \in G$ and for all $i = 1, \dots, m$ there exists a $j \in \{1, \dots, m\}$ such that $gV_i \subseteq V_j$);
3. **a field extension subgroup**, or of **class** $\mathcal{C}_3$, if there exist a prime s dividing $2n$, a structure of $\mathbb{F}_{\ell^s}$-vector space on $\mathbb{F}_\ell^{2n}$, and a subgroup H of G of index s such that H acts on $\mathbb{F}_\ell^{2n}$ preserving the $\mathbb{F}_{\ell^s}$-structure;
4. a **tensor product subgroup**, or of **class** $\mathcal{C}_4$, if there is a decomposition $\mathbb{F}_\ell^{2n} \cong V_1 \otimes V_2$ (where V_1, V_2 are $\mathbb{F}_\ell$-vector spaces) and for each $g \in G$ there exist $g_1 \in \mathrm{GL}(V_1)$ and $g_2 \in \mathrm{GL}(V_2)$ for which $g = g_1 \otimes g_2$.

We shall also have to deal with the exceptional class $\mathcal{S}$:

Definition 4.3.13. (cf. [19, Definition 2.1.3]) A subgroup H of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ is said to be of class $\mathcal{S}$ if and only if all of the following hold:

1. $\mathbb{P}H$ is almost simple;
2. H does not contain $\mathrm{Sp}_{2n}(\mathbb{F}_\ell)$;
3. H^∞ acts absolutely irreducibly on $\mathbb{F}_\ell^{2n}$.

A general philosophy (cf. for example [118], especially §3, or [24, Remark 2.1]) predicts that groups in class $\mathcal{S}$ should come in two different flavours. On one hand, there should exist finitely many groups $G_1, \dots, G_k$ that embed in $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ for infinite families of primes ℓ; we shall refer to these as **constant groups**. On the other hand, if $\mathcal{G}$ is an algebraic group over $\mathbb{Z}$ admitting an irreducible, symplectic representation of dimension $2n$, then the corresponding embedding $\mathcal{G} \hookrightarrow \mathrm{GSp}_{2n,\mathbb{Z}}$ should give rise – for almost all primes ℓ – to a maximal subgroup $\mathcal{G}(\mathbb{F}_\ell)$ of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. We shall refer to groups arising in this way as **groups of Lie type**. We do not turn these notions into precise definitions, but it will be clear from sections 4.6 and 4.10 that there are indeed two different kinds of class-$\mathcal{S}$ subgroups, and that they need to be treated in different ways.

We are now finally ready to state the following classification theorem, essentially due to Aschbacher (but see also [49] and [19]):

Theorem 4.3.14. *(Aschbacher [5]) Let n be an odd integer, ℓ be an odd prime, and G be a maximal proper subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ not containing $\mathrm{Sp}_{2n}(\mathbb{F}_\ell)$. Then one of the following holds:*

1. *G is of class $\mathcal{C}_1$;*
2. *G is of class $\mathcal{C}_2$, stabilizing an m-decomposition for some $m \geq 2$ dividing $2n$;*
3. *G is of class $\mathcal{C}_3$ for some prime s dividing $2n$;*
4. *G is of class $\mathcal{C}_4$, and more precisely G is isomorphic to $\mathrm{GSp}_{2m}(\mathbb{F}_\ell) \otimes \mathrm{CGO}_t(\mathbb{F}_\ell)$, where m and $t \geq 3$ are integers such that $2mt = 2n$ (we call (m,t) the **type** of G);*
5. *G is of class $\mathcal{S}$.*

The proof of theorem 4.1.3 essentially consists in going through the list provided by theorem 4.3.14 to show that, for ℓ large enough, G_ℓ cannot be contained in any proper maximal subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$, and therefore the equality $G_\ell = \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ must hold.

4.4 Reducible, imprimitive and field extension cases

Recall from the introduction that we denote by A/K an abelian variety of dimension g with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, and by G_ℓ the image of the representation $\rho_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\, A[\ell]$. At least for $\ell > b(A/K)$ we know from corollary 4.2.4 that $G_\ell \subseteq \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. Suppose now that G_ℓ does not contain $\mathrm{Sp}_{2g}(\mathbb{F}_\ell)$: then G_ℓ is contained in one of the maximal subgroups listed in theorem 4.3.14. The following proposition shows that cases 1 through 3 of that theorem cannot arise for ℓ large enough:

Proposition 4.4.1. *Let G be a maximal proper subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. Suppose G is*

1. *reducible: then $\ell \leq b_0(A/K)$.*
2. *imprimitive: then $\ell \leq b_0(A/K; g!)$.*
3. *a field extension subgroup: then $\ell \leq b_0(A^2/K; g)^{1/2g}$.*

Proof. Replacing K with an extension of degree at most $g!$ or g in cases 2 and 3, we can assume that G_ℓ stabilizes a subspace (cases 1 and 2), or that its centralizer is strictly larger than $\mathbb{F}_\ell^\times$ (case 3). The claim then follows from lemmas 3.3.17 and 3.3.18. □

4.5 Groups of Lie type with socle $\mathrm{PSL}_2(\mathbb{F}_\ell)$

We now consider maximal class-$\mathcal{S}$ subgroups G of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ that satisfy $\mathrm{soc}(\mathbb{P}G) \cong \mathrm{PSL}_2(\mathbb{F}_\ell)$. There are two reasons why we single out this case: on one hand, it is not hard to construct (for all n and most ℓ) an explicit family of maximal subgroups of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ having this shape, so this is clearly a case we need to treat; on the other hand, as we shall show in section 4.10, for most values of n this is in fact the *only* kind of class-$\mathcal{S}$ subgroup of Lie type of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$.
To see how such subgroups with socle $\mathrm{PSL}_2(\mathbb{F}_\ell)$ arise, denote by $V_1 := \mathbb{F}_\ell^2$ the defininining representation of either $\mathrm{GL}_2(\mathbb{F}_\ell)$ or $\mathrm{SL}_2(\mathbb{F}_\ell)$, and consider, for every positive integer n, the $(2n-1)$-th symmetric power of V_1, which we denote by V_{2n-1}; it is a symplectic representation of $\mathrm{GL}_2(\mathbb{F}_\ell)$

or $\mathrm{SL}_2(\mathbb{F}_\ell)$ respectively. Moreover, for $\ell > 2n$ this representation is absolutely irreducible ([19, Proposition 5.3.6 (i)]), hence its image gives rise to a maximal class-$\mathcal{S}$ subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ with socle $\mathrm{PSL}_2(\mathbb{F}_\ell)$. We denote by

$$\sigma_{2n-1} : \mathrm{GL}_2(\mathbb{F}_\ell) \to \mathrm{GSp}(V_{2n-1}) \cong \mathrm{GSp}_{2n}(\mathbb{F}_\ell)$$

the representation thus obtained, and by S_{2n-1} the image of $\mathrm{GL}_2(\mathbb{F}_\ell)$ in $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$. As the following lemma shows, the group S_{2n-1} is the only one we need to consider:

Lemma 4.5.1. *([19, Proposition 5.3.6 (i)]) Let $\ell > 2n$ be a prime number and let G be a maximal class-$\mathcal{S}$ subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ such that $\operatorname{soc} \mathbb{P}G \cong \mathrm{PSL}_2(\mathbb{F}_\ell)$. Then (up to conjugation in $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$) we have $\mathbb{P}G = \mathbb{P}S_{2n-1}$.*

We now turn to the application to abelian varieties. Suppose once more that A/K is an abelian variety of dimension g with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, and suppose that for some prime $\ell > 2g$ the group G_ℓ is contained in a maximal class-$\mathcal{S}$ subgroup G of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ with projective socle $\mathrm{PSL}_2(\mathbb{F}_\ell)$. By the previous lemma, we can assume $\mathbb{P}G = \mathbb{P}S_{2g-1}$. In this situation, the assumption $G_\ell \subseteq G$ implies that for every $h \in G_\ell$ there exist a scalar $\lambda \in \mathbb{F}_\ell^\times$ and an element $M \in \mathrm{GL}_2(\mathbb{F}_\ell)$ such that $h = \lambda \cdot \sigma_{2g-1}(M)$. In particular, the eigenvalues of h are given by the (multi)set

$$\left\{ \lambda \mu^j \nu^{2g-1-j} \;\middle|\; j = 0, \dots, 2g-1 \right\}, \tag{4.1}$$

where μ, ν are the eigenvalues of M. Notice that the eigenvalues of M lie either in $\mathbb{F}_\ell$ or in its (unique) quadratic extension, hence all eigenvalues of h are elements of $\mathbb{F}_{\ell^2}$. We shall now show that (for ℓ large enough) this description of the eigenvalues of h contradicts what is known about the representation ρ_ℓ restricted to the inertia at ℓ. More precisely, let $\mathfrak{l}$ be a place of K above the prime ℓ, let $I_\mathfrak{l} \subseteq \mathrm{Gal}(\overline{K}/K)$ be the inertia group at $\mathfrak{l}$, and write $I_\mathfrak{l}^t$ for the *tame* inertia group at $\mathfrak{l}$. Under a semistability hypothesis, the action of $I_\mathfrak{l}$ on $A[\ell]$ factors through $I_\mathfrak{l}^t$, and is described by the following theorem of Raynaud:

Theorem 4.5.2. *([104, Corollaire 3.4.4]) Suppose A has semistable reduction at $\mathfrak{l}$: then the wild inertia subgroup of $I_\mathfrak{l}$ acts trivially on $A[\ell]$, so the action of $I_\mathfrak{l}$ factors through $I_\mathfrak{l}^t$. Let V be a Jordan-Hölder quotient of $A[\ell]$ for the action of $I_\mathfrak{l}^t$. Suppose V is of dimension n over $\mathbb{F}_\ell$, and let e be the ramification index of $\mathfrak{l}$ over ℓ. There exist integers $e_1, \dots, e_n$ such that:*

- *V has a structure of $\mathbb{F}_{\ell^n}$-vector space;*
- *the action of $I_\mathfrak{l}^t$ on V is given by a character $\psi : I_\mathfrak{l}^t \to \mathbb{F}_{\ell^n}^\times$;*
- *$\psi = \varphi_1^{e_1} \dots \varphi_n^{e_n}$, where $\varphi_1, \dots, \varphi_n$ are the fundamental characters of $I_\mathfrak{l}^t$ of level n;*
- *for every $i = 1, \dots, n$ the inequality $0 \le e_i \le e$ holds.*

Remark 4.5.3. Raynaud's theorem is usually stated for places of *good* reduction. However, as it was shown in [61, Lemma 4.9], the extension to the semistable case follows easily upon applying results of Grothendieck [30].

Remark 4.5.4. By construction the fundamental characters of level n are *surjective* morphisms $I_\mathfrak{l}^t \to \mathbb{F}_{\ell^n}^\times$. Moreover, the norm of a fundamental character of level n (taken from $\mathbb{F}_{\ell^n}$ to $\mathbb{F}_\ell$) is the unique fundamental character of level 1. If furthermore $\mathfrak{l}$ is unramified in K, then this unique character of level 1 is χ_ℓ, the cyclotomic character mod ℓ.

Notation. For the rest of this section we suppose that ℓ is a prime for which there exists a place $\mathfrak{l}$ of K of characteristic ℓ at which A has either good or bad semistable reduction.

Let now $W_1, \dots, W_k$ be the sequence of Jordan-Hölder quotients of $A[\ell]$ under the action of $I_\mathfrak{l}^t$, and $\psi_1, \dots, \psi_k$ be the corresponding characters as in Raynaud's theorem. Also write $n_i = \dim W_i$ and suppose, for the rest of the section, that ℓ is unramified in K.

Lemma 4.5.5. *Every n_i is at most 2.*

Proof. Let W be any simple Jordan-Hölder quotient of $A[\ell]$ and let ψ be the associated character. Suppose that the image of ψ is contained in $\mathbb{F}_{\ell^k}^\times$ for a certain $k \geq 1$, and let σ be a generator of $\mathrm{Gal}\,(\mathbb{F}_{\ell^k}/\mathbb{F}_\ell)$. Since the action of $I_\mathfrak{l}^t$ on W can be diagonalized over $\mathbb{F}_{\ell^k}$, we can find a vector $v \in W \otimes_{\mathbb{F}_\ell} \mathbb{F}_{\ell^k}$ that is a common eigenvector for the action of $I_\mathfrak{l}^t$. The $\mathbb{F}_\ell$-vector subspace of $W \otimes_{\mathbb{F}_\ell} \mathbb{F}_{\ell^k}$ spanned by $v, \sigma v, \dots, \sigma^{k-1} v$ is by construction σ-stable, hence it descends to a $\mathbb{F}_\ell$-subspace W' of W, and it is clear by construction that W' is also stable under the action of $I_\mathfrak{l}^t$. As W is irreducible and W' is nontrivial we must have $W' = W$, and since $\dim W' \leq k$ we have $\dim W \leq k$. In our situation, we have already remarked that all the eigenvalues of every element of G_ℓ lie in $\mathbb{F}_{\ell^2}$, hence in particular the same is true for the eigenvalues of the action of $I_\mathfrak{l}^t$. It follows that the image of ψ is entirely contained in $\mathbb{F}_{\ell^2}$, and the previous argument shows that W is of dimension at most 2. □

In view of Raynaud's theorem and of the previous lemma, the only characters through which $I_\mathfrak{l}^t$ can act on $A[\ell]$ are the fundamental characters of level 1 and 2, along with the trivial character. Denote by m_0 (resp. m_1, m_2) the number of Jordan-Hölder quotients of $A[\ell]$ on which $I_\mathfrak{l}^t$ acts trivially (resp. through χ_ℓ, through one of the fundamental characters of level 2). As $A[\ell]$ is of dimension $2g$, the dimensions of its simple Jordan-Hölder quotients must add up to $2g$, and so we have

$$m_0 + m_1 + 2m_2 = 2g. \tag{4.2}$$

These three numbers also satisfy another numerical relation:

Lemma 4.5.6. *Suppose $\ell > g + 1$ is unramified in K: then $m_0 = m_1$.*

Proof. Notice that since ℓ is unramified in K the exponents e_i in Raynaud's theorem are all either 0 or 1. Write $\varphi_1, \varphi_2 = \varphi_1^\ell$ for the two fundamental characters of level 2. If W is a simple Jordan-Hölder quotient of $A[\ell]$ of dimension 2, the action of $x \in I_\mathfrak{l}^t$ on W has eigenvalues $\varphi_1(x)$ and $\varphi_2(x)$, hence its determinant is $\varphi_1(x)\varphi_2(x) = \chi_\ell(x)$. On the other hand, the determinant of the action on 1-dimensional simple quotients is either 1 (if the action is trivial) or $\chi_\ell(x)$ (if the action is through χ_ℓ). It follows that

$$\chi_\ell(x)^g = \det\left(\rho_\ell(x) : A[\ell] \to A[\ell]\right) = \prod_{W_i} \det\left(\rho_\ell(x) : W_i \to W_i\right) = \chi_\ell(x)^{m_1} \chi_\ell(x)^{m_2} \quad \forall x \in I_\mathfrak{l}^t,$$

i.e. $\chi_\ell^{m_1+m_2-g} \equiv 1$ on $I_\mathfrak{l}^t$. Since ℓ is unramified in K, the order of the image of χ_ℓ is $\ell-1$, hence we must have $(\ell-1) \mid m_1+m_2-g$. However, $|m_1+m_2-g| \leq g$ by equation (4.2), and since $\ell-1 > g$ by assumption the only possibility is $m_1+m_2=g$. Together with $m_0+m_1+2m_2=2g$ this yields $m_0=m_1$ as claimed. □

The next step is to show that in fact there are no inertia invariants if ℓ is sufficiently large with respect to g:

Lemma 4.5.7. *Suppose $g \geq 3$. If $\ell > g(2g-1)+1$ is unramified in K, then $m_0 = 0$.*

Proof. The previous lemmas imply that $m_1+m_2 = g \geq 3$, hence we have $\max\{m_1,m_2\} \geq 2$. Suppose by contradiction that $m_0 \geq 1$. By definition of m_0, m_1 and m_2, for every $x \in I_\mathfrak{l}^t$ the eigenvalues of $\rho_\ell(x)$ are $\{1, \chi_\ell(x), \varphi_1(x), \varphi_1(x)^\ell\}$, with multiplicities given respectively by m_0, m_1, m_2 and m_2. On the other hand, we know from (4.1) that the eigenvalues of $\rho_\ell(x)$ can be written as $\{\lambda\mu^j\nu^{2g-1-j} \mid j=0,\dots,2g-1\}$ for some $\lambda \in \mathbb{F}_\ell^\times$ and $\mu,\nu \in \mathbb{F}_{\ell^2}^\times$. Now for all $x \in I_\mathfrak{l}^t$ the operator $\rho_\ell(x)$ admits an eigenvalue of multiplicity at least 2 (since $\max\{m_1,m_2\} \geq 2$) and it also has 1 among its eigenvalues (since $m_0 \geq 1$): thus there exist two indices $0 \leq j_1 < j_2 \leq 2g-1$ (depending on x) such that $\lambda\mu^{j_1}\nu^{2g-1-j_1} = \lambda\mu^{j_2}\nu^{2g-1-j_2}$, and an index $0 \leq j_3 \leq 2g-1$ (depending on x, and not necessarily distinct from j_1, j_2) such that $\lambda\mu^{j_3}\nu^{2g-1-j_3} = 1$. These equations can be rewritten as

$$\begin{cases} (\mu/\nu)^{j_1-j_2} = 1 \\ \lambda = \mu^{-j_3}\nu^{j_3-2g+1} = (\mu/\nu)^{-j_3}\nu^{1-2g}. \end{cases}$$

On the other hand, the fact that $\det\rho_\ell(x) = \chi_\ell(x)^g$ yields

$$\chi_\ell(x)^g = \det\rho_\ell(x) = \prod_{j=0}^{2g-1}\left(\lambda\mu^j\nu^{2n-1-j}\right) = \lambda^{2g}(\mu\nu)^{2g^2-g}, \tag{4.3}$$

and upon replacing λ by $(\mu/\nu)^{-j_3}\nu^{1-2g}$ we get $\chi_\ell(x)^g = (\mu/\nu)^{g(2g-1-2j_3)}$. Finally, raising both sides of this equation to the (j_1-j_2)-th power and using $(\mu/\nu)^{j_1-j_2}=1$ we find

$$\chi_\ell(x)^{g(j_1-j_2)} = (\mu/\nu)^{g(j_1-j_2)(2g-1-2j_3)} = 1,$$

which proves in particular that $\operatorname{ord}\chi_\ell(x) \leq g(j_2-j_1) \leq g(2g-1)$ for all $x \in I_\mathfrak{l}^t$. But since ℓ is unramified in K the image of $\chi_\ell|_{I_\mathfrak{l}^t}$ is a cyclic group of order $\ell-1 > g(2g-1)$: taking an $x \in I_\mathfrak{l}^t$ such that $\chi_\ell(x)$ generates $\chi_\ell(I_\mathfrak{l}^t)$ gives a contradiction, which shows that we must in fact have $m_0=0$. □

We have thus proved that for $\ell > g(2g-1)+1$ we necessarily have $m_0=m_1=0$ and $m_2=g$. It remains to show that this is impossible as well:

Lemma 4.5.8. *Suppose $\ell \geq 2g$ is unramified in K: then we cannot have $m_2 = g$.*

Proof. The proof is very similar to that of the previous lemma, so we keep the same notation. Let x be any element of $I_\mathfrak{l}^t$. The assumption $m_2=g$ implies (by an obvious pigeonhole argument) that we can find two indices $0 \leq j_1 < j_2 \leq 2g-1$ such that $j_2-j_1 \leq 2$ and $\lambda\mu^{j_1}\nu^{2g-1-j_1} = \lambda\mu^{j_2}\nu^{2g-1-j_2}$, which implies $(\mu/\nu)^{j_2-j_1}=1$ and therefore $\mu/\nu = \pm 1$. Moreover there exists an index $0 \leq j \leq 2g-1$ such that $\lambda\mu^j\nu^{2g-1-j} = \varphi_1(x)$, hence $\lambda^{2g} = \varphi_1(x)^{2g}\nu^{2g(1-2g)}$. Equation (4.3) now implies

$$\chi_\ell(x)^g = \lambda^{2g}(\mu\nu)^{g(2g-1)} = \varphi_1(x)^{2g}(\mu/\nu)^{g(2g-1)} = \pm\varphi_1(x)^{2g},$$

whence, using $\chi_\ell(x) = \varphi_1(x)\varphi_2(x) = \varphi_1(x)^{\ell+1}$, we see that $\varphi_1(x)^{g(\ell-1)} = \pm 1$ for all $x \in I_\mathfrak{l}^t$. This implies that the cyclic group $\varphi_1(I_\mathfrak{l}^t)$ has order at most $2g(\ell-1)$, but on the other hand (since ℓ is unramified in K) we know that $\left|\varphi_1(I_\mathfrak{l}^t)\right| = \ell^2 - 1$. This implies $\ell + 1 \leq 2g$, contrary to our assumptions. □

Putting together the last three lemmas we have

Proposition 4.5.9. *Suppose $\ell > 2g(g-1)+1$ is a prime unramified in K and such that there is at least one place $\mathfrak{l}$ of K of characteristic ℓ at which A has semistable reduction. Then G_ℓ cannot be contained in a maximal class-$\mathcal{S}$ subgroup G of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ with $\operatorname{soc} \mathbb{P}G \cong \mathrm{PSL}_2(\mathbb{F}_\ell)$.*

4.6 Constant groups in class $\mathcal{S}$

The analysis of the constant subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ is greatly simplified by the following theorems of Larsen-Pink and Collins:

Theorem 4.6.1. *(Larsen-Pink [60, Theorem 0.2]) For every positive integer n there exists a constant $J'(n)$ with the following property: any finite subgroup Γ of $\mathrm{GL}_n(k)$ over any field k possesses normal subgroups $\Gamma_3 \subset \Gamma_2 \subset \Gamma_1$ such that*

(a) $[\Gamma : \Gamma_1] \leq J'(n)$;

(b) either $\Gamma_1 = \Gamma_2$, *or* $p := \mathrm{char}(k)$ *is positive and* Γ_1/Γ_2 *is a direct product of finite simple groups of Lie type in characteristic* p;

(c) Γ_2/Γ_3 *is abelian, of order not divisible by* $\mathrm{char}(k)$;

(d) either $\Gamma_3 = \{1\}$, *or* $p := \mathrm{char}(k)$ *is positive and* Γ_3 *is a p-group.*

Theorem 4.6.2. *([22, Theorem A]) One can take* $J'(n) := \begin{cases} (n+2)!, & \text{if } n \geq 71 \\ n^4(n+2)!, & \text{if } n < 71 \end{cases}$, *which is also optimal for* $n \geq 71$. *Furthermore, if in the previous theorem we restrict to fields k such that* $\mathrm{char}\, k \nmid (n+1)(n+2)$, *then one can replace $J'(n)$ by* $J(n) := \begin{cases} (n+1)!, & \text{if } n \geq 71 \\ n^4(n+2)!, & \text{if } n < 71 \end{cases}$

Remark 4.6.3. Collin's theorem is in fact more precise and gives the optimal value of $J'(n)$ also for $n \leq 71$. Using this improved bound would not change our final result (theorem 4.1.3), and we have therefore chosen to use the simpler expression given above.

Theorem 4.6.1 immediately implies:

Proposition 4.6.4. *Let ℓ, g be such that $\ell \nmid (2g+1)(2g+2)$. Suppose $G \subseteq \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ is a maximal subgroup of class $\mathcal{S}$ and satisfies $|\mathbb{P}G| > J(2g)$: then the socle of $\mathbb{P}G$ is a simple group of Lie type in characteristic ℓ.*

Proof. Apply theorem 4.6.1 to G. Notice first that Γ_3 is trivial: indeed, Γ_3 is a solvable normal subgroup of G, so $\mathbb{P}\Gamma_3$ is a solvable normal subgroup of $\mathbb{P}G$, which is almost-simple since G is of class $\mathcal{S}$. It follows from lemma 4.3.3 that $\mathbb{P}\Gamma_3$ is trivial, so Γ_3 is a subgroup of the group of homotheties in $\mathrm{GL}_{2g}(\mathbb{F}_\ell)$, which has order prime to ℓ, hence $\Gamma_3 = \{1\}$ as claimed. The same argument now shows that $\Gamma_2 \subseteq \mathbb{F}_\ell^\times \cdot \mathrm{Id}$, for otherwise $\mathbb{P}\Gamma_2$ would be an abelian (in particular solvable) normal subgroup of $\mathbb{P}G$. This implies in particular that Γ_1 and Γ_2 commute, and that $\mathbb{P}(\Gamma_1\Gamma_2) = \mathbb{P}\Gamma_1$. Notice that $\mathbb{P}\Gamma_1$ cannot be trivial, for otherwise we would have $|\mathbb{P}G| \leq J(2n)|\mathbb{P}(\Gamma_1)| = J(2n)$, contradicting the hypothesis; hence $\mathbb{P}\Gamma_1$ is a nontrivial normal subgroup of $\mathbb{P}G$, so it contains $\mathrm{soc}(G)$. On the other hand, the fact that Γ_2 consists entirely of homotheties implies that $\mathbb{P}\Gamma_1$ is a quotient of Γ_1/Γ_2, hence in particular a direct product of finite simple groups of Lie type in characteristic ℓ. Lemma 4.3.7 now implies that $\operatorname{soc}\mathbb{P}G = (\Gamma_1/\Gamma_2)^\infty$ is of Lie type in characteristic ℓ. □

Proposition 4.6.5. *Let ℓ be a prime such that there is a place $\mathfrak{l}$ of K of residual characteristic ℓ at which A has either good or bad semistable reduction. If ℓ is unramified in K and not less than $g+2$, then $|\mathbb{P}G_\ell| \geq \ell - 1$.*

Proof. We take the notation of section 4.5; in particular, we let $W_1, \ldots, W_k$ be the simple Jordan-Hölder quotients of $A[\ell]$ under the action of the inertia group $I_\mathfrak{l}$ (or equivalently, of the tame inertia group $I_\mathfrak{l}^t$), and $\psi_1, \ldots, \psi_k$ be the characters associated with the W_i's by Raynaud's theorem 4.5.2. Let N be the order of $|\mathbb{P}G_\ell|$, and notice that for every $y \in G_\ell$ the projective image of y^N is trivial, that is, y^N is a multiple of the identity, and in particular has a unique eigenvalue of multiplicity $2g$. Since for $x \in I_\mathfrak{l}^t$ the eigenvalues of $\rho_\ell(x)$ are given by the Galois conjugates of the various $\psi_i(x)$, this implies that for all $i, j = 1, \ldots, k$, for all integers $t \geq 0$, and for all $x \in I_\mathfrak{l}$ we have

$$\psi_i(x)^{\ell^t N} = \psi_j(x)^N. \tag{4.4}$$

We now distinguish three cases:

1. At least one of the W_i's is of dimension ≥ 2: without loss of generality, we can assume that $n := \dim W_1$ is at least 2. Let ψ be the associated character. By Raynaud's theorem, there are integers $e_0, \ldots, e_{n-1} \in \{0, 1\}$ such that $\psi = \varphi^{\sum_{i=0}^{n-1} e_i \ell^i}$, where φ is a fundamental character of level n. Note that we cannot have $e_i = 1$ for $i = 0, \ldots, n-1$, for otherwise we would have $\psi = \chi_\ell$, which contradicts the fact that W_1 is of dimension $n > 1$ (cf. the proof of lemma 4.5.5). In particular, since for all integers $t \geq 0$ the character φ^{ℓ^t} is a Galois conjugate of φ, replacing φ with φ^{ℓ^t} for a suitable t we can assume that $e_{n-1} = 0$ (notice that replacing φ with φ^ℓ has the effect of permuting cyclically the integers e_i, at least one of which is zero). Now φ has exact order $\ell^n - 1$, so $\psi = \varphi^{\sum_{i=0}^{n-1} e_i \ell^i}$ has order at least
$$\frac{\ell^n - 1}{\sum_{i=0}^{n-1} e_i \ell^i} \geq \frac{\ell^n - 1}{\sum_{i=0}^{n-2} \ell^i} = \frac{(\ell^n - 1)(\ell - 1)}{(\ell^{n-1} - 1)} \geq \ell(\ell - 1),$$
that is to say, there is an $x \in I_\mathfrak{l}^t$ such that $\psi(x)$ has order at least $\ell(\ell-1)$. Consider now equation (4.4) for this specific x, for $\psi_i = \psi_j = \psi$ and for $t = 1$: it gives $\psi(x)^{(\ell-1)\cdot N} = 1$, so $\psi(x)$ has order at most $(\ell-1)\cdot N$. Thus we obtain $(\ell-1)\cdot N \geq \ell(\ell-1)$, that is $N \geq \ell > \ell - 1$ as claimed.

2. All the W_i's are of dimension 1, for at least one index i we have $\psi_i = 1$, and for at least one index j we have $\psi_j = \chi_\ell$: then for all $x \in I_\mathfrak{l}^t$ we have $\psi_j(x)^N = \psi_i(x)^N$, that is, $\chi_\ell(x)^N = 1$ for all $x \in I_\mathfrak{l}^t$. As χ_ℓ has exact order $\ell - 1$, this implies $N \geq \ell - 1$.

3. All the W_i's are of dimension 1 and all characters ψ_i are equal to each other (and in particular are either all trivial or all equal to the cyclotomic character χ_ℓ): in this case there are exactly $k = 2g$ simple Jordan-Hölder quotients, and from the equality
$$\chi_\ell(x)^g = \det \rho_\ell(x) = \prod_{i=1}^{2g} \psi_i(x) = \begin{cases} 1, \text{ if } \psi_i = 1 \text{ for every } i \\ \chi_\ell(x)^{2g}, \text{ if } \psi_i = \chi_\ell \text{ for every } i \end{cases}$$
we find $\chi_\ell(x)^g = 1$ for all $x \in I_\mathfrak{l}^t$, which contradicts the fact that the order of χ_ℓ is $\ell - 1 > g$.

□

Corollary 4.6.6. *Let $\ell \geq J(2g) + 2$ be a prime unramified in K. Suppose that there exists a place $\mathfrak{l}$ of K, of residual characteristic ℓ, at which A has semistable reduction: then $|\mathbb{P}G_\ell| > J(2g)$.*

Remark 4.6.7. Proposition 4.6.4 should be interpreted as saying that the order of the constant groups appearing as maximal subgroups of $\mathbb{P}\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ is bounded by $J(2g)$ (for large enough g, equality is attained by the natural $2g$-dimensional representation of S_{2g+1}). Corollary 4.6.6 then amounts to saying that for $\ell > J(2g) + 1$ (and under a suitable semistability hypothesis) the action of Galois cannot factor through a constant group of class $\mathcal{S}$.

4.7 The tensor product case I

We are now left with the task of showing that, for ℓ large enough, the group G_ℓ cannot be contained in a tensor product subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$. Let us briefly explain the key idea behind the proof, which goes back to Serre (cf. [120]). If G_ℓ is contained in a tensor product subgroup, this forces the eigenvalues of any $x \in G_\ell$ to satisfy a number of additional multiplicative relations that do not hold for a sufficiently generic element of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$: we will therefore be able to show that G_ℓ is not contained in a tensor product subgroup as soon as we have at our disposal an element of G_ℓ whose eigenvalues do not satisfy any multiplicative relations except for the "obvious" ones. We shall look for such an element among those of the form $\rho_\ell(\mathrm{Fr}_v)$, where Fr_v is a Frobenius element associated with a place v of K: since the eigenvalues of $\rho_\ell(\mathrm{Fr}_v)$ are independent of ℓ, if for a certain prime ℓ_0 the eigenvalues of $\rho_{\ell_0}(\mathrm{Fr}_v)$ do not satisfy these additional relations, then the same is true for the eigenvalues of $\rho_\ell(\mathrm{Fr}_v)$ for all but finitely many primes ℓ. This will be enough to conclude that, for ℓ large enough, G_ℓ is not contained in a tensor product subgroup.
We split the analysis of tensor product subgroups in two parts: in the present section we show that, given such a "generic" Frobenius element, we can indeed give an explicit bound on the largest prime ℓ for which G_ℓ can be contained in a tensor product subgroup; then, in section 4.9, we shall show how, when $g = 3$, Chebotarev's density theorem enables us to find a suitable Frobenius element.

To carry out both parts of this program we shall need to study Frobenius elements and their eigenvalues in some detail. We let Ω_K denote the set of finite places of K, and for each $v \in \Omega_K$

we write p_v for the residual characteristic and q_v for the cardinality of the residue field at v. We also write $\mathrm{Fr}_v \in \mathrm{Gal}(\overline{K}/K)$ for a Frobenius element at v. If v is a place of K of good reduction for A, the characteristic polynomial of $\rho_{\ell^\infty}(\mathrm{Fr}_v)$ does not depend on ℓ (as long as $v \nmid \ell$), and will be denoted by $f_v(x) \in \mathbb{Z}[x]$. We shall write $\mu_1, \dots, \mu_{2g}$ for the roots of $f_v(x)$ in $\overline{\mathbb{Q}}$, and call these algebraic integers the **eigenvalues of** Fr_v.
The splitting field of $f_v(x)$ is a Galois extension of $\mathbb{Q}$ which we call $F(v)$. If ℓ is a prime not lying below v, let $\mathfrak{l}$ be any prime of $F(v)$ lying above ℓ, and let $\mathbb{F}_\mathfrak{l}$ be the residue field at $\mathfrak{l}$. Since the μ_i's are algebraic *integers*, it makes sense to consider their reductions modulo $\mathfrak{l}$, which are elements of $\overline{\mathbb{F}_\ell}^\times$ which we will denote by $\overline{\mu_1}, \dots, \overline{\mu_{2g}}$; clearly these $\overline{\mu_i}$'s can also be identified with the roots in $\overline{\mathbb{F}_\ell}$ of the characteristic polynomial of $\rho_\ell(\mathrm{Fr}_v)$. When speaking of the roots $\overline{\mu_1}, \dots, \overline{\mu_{2g}}$ of the characteristic polynomial of $\rho_\ell(\mathrm{Fr}_v)$ we shall always implicitly assume that this identification has been made.

Lemma 4.7.1. *The splitting field $F(v)$ of the characteristic polynomial $f_v(x)$ of Fr_v has Galois group isomorphic to a subgroup of $(\mathbb{Z}/2\mathbb{Z})^g \rtimes S_g$, so it has degree at most $2^g g!$ over $\mathbb{Q}$.*

Proof. Immediate from the relation $x^{2g} f_v\left(q_v x^{-1}\right) = q_v^g f_v(x)$, which in turn follows from $\rho_{\ell^\infty}(\mathrm{Fr}_v)$ being an element of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ for any sufficiently large prime ℓ and from the Weil conjectures. □

We shall need the following basic facts from group theory, whose proof is completely straightforward:

Lemma 4.7.2. *Let m, n be positive integers.*

1. *Let $\ell \geq 3$ be a prime. The groups $\mathrm{Sp}_{2m}(\mathbb{F}_\ell) \otimes \mathrm{SO}_{2n+1}(\mathbb{F}_\ell)$ and $\mathrm{Sp}_{2m}(\mathbb{F}_\ell) \otimes \mathrm{GO}_{2n+1}(\mathbb{F}_\ell)$ coincide.*
2. *Let F be a field not of characteristic 2 and h be an element of $\mathrm{SO}_{2n+1}(F)$. The multiset Ψ of eigenvalues of h can be written as $\left\{\beta_1, \dots, \beta_n, 1, \beta_1^{-1}, \dots, \beta_n^{-1}\right\}$ for certain $\beta_1, \dots, \beta_n \in \overline{F}^\times$.*
3. *Suppose m, n are odd and let $g = mn$. Let G be a maximal subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ of tensor product type (m, n), that is, $G \cong \mathrm{GSp}_{2m}(\mathbb{F}_\ell) \otimes \mathrm{CGO}_n(\mathbb{F}_\ell)$. For every $h \in G$, the eigenvalues of h can be written as $\left\{\lambda_i \beta_j, \lambda_i, \lambda_i \beta_j^{-1} \;\middle|\; i = 1, \dots, 2m,\ j = 1, \dots, \frac{n-1}{2}\right\}$ for certain $\lambda_1, \dots, \lambda_{2m}, \beta_1, \dots, \beta_{\frac{n-1}{2}}$ in $\overline{\mathbb{F}_\ell}^\times$.*

We now start investigating the multiplicative relations satisfied by the eigenvalues of an operator lying in a tensor product subgroup. Even though in general there may be additional relations, by part (3) of the previous lemma we already know a large number of equations these eigenvalues must satisfy; to state them more concisely, we introduce the following definition:

Definition 4.7.3. We let V_{mn} be the affine scheme cut in $\mathbb{A}_\mathbb{Z}^{2g}$ (with variables $z_1, \dots, z_{2m}$ and x_{ij}, y_{ij} for $i = 1, \dots, 2m$ and $j = 1, \dots, \frac{n-1}{2}$) by the equations

$$\begin{cases} x_{ij} y_{ij} = z_i^2 & \text{for } i = 1, \dots, 2m \text{ and } j = 1, \dots, \frac{n-1}{2} \\ z_k x_{ij} = z_i x_{kj} & \text{for } i, k = 1, \dots, 2m \text{ and } j = 1, \dots, \frac{n-1}{2} \\ z_k y_{ij} = z_i y_{kj} & \text{for } i, k = 1, \dots, 2m \text{ and } j = 1, \dots, \frac{n-1}{2} \end{cases}$$

We denote $v = (z_k, x_{ij}, y_{ij})$ a point in $\mathbb{A}_\mathbb{Z}^{2g}$ and let elements $\sigma \in S_{2g}$ act on $\mathbb{A}_\mathbb{Z}^{2g}$ by permuting the coordinates in the natural way. For every $\sigma \in S_{2g}$ we also consider the scheme V_{mn}^σ defined

by $\left\{ v \in \mathbb{A}_{\mathbb{Z}}^{2g} \;\middle|\; \sigma(v) \in V_{mn} \right\}$. We also let $\mathcal{P}_{mn}^{\sigma}$ be a set of homogeneous binomials of degree 2 with coefficients in $\{\pm 1\}$ that generate the ideal of V_{mn}^{σ}: it is clear by the definition of V that such polynomials exist. Finally, we let U_{mn}^{σ} be the open subscheme of V_{mn}^{σ} over which $\prod_{i=1}^{m} z_i \prod_{j=1}^{(n-1)/2} x_{ij} y_{ij}$ is invertible (note that this condition is invariant under the action of S_{2g}), and to ease the notation we set $U_{mn} := U_{mn}^{\mathrm{id}}$.

Lemma 4.7.4. *Let F be a field. For a $2g$-tuple $(w_1, \dots, w_{2g})$ of elements of $F^{\times}$ the following are equivalent:*

1. *there exists a permutation $\sigma \in S_{2g}$ such that $(w_1, \dots, w_{2g}) \in U_{mn}^{\sigma}(F)$;*
2. *there exist $\lambda_1, \dots, \lambda_{2m}, \beta_1, \dots, \beta_{\frac{n-1}{2}} \in F^{\times}$ such that $w_1, \dots, w_{2g}$ equal (in some order) the $2g$ numbers $\lambda_i, \lambda_i \beta_j^{\pm 1}$ for $i = 1, \dots, 2m$ and $j = 1, \dots, \frac{n-1}{2}$.*

Proof. Notice that both conditions are invariant under the action of S_{2g}, so we consider the statement up to permutation of the coordinates. Assume first that (2) holds: then we obtain a point of $U_{mn}(F)$ by setting, for $i = 1, \dots, 2m$ and $j = 1, \dots, \frac{n-1}{2}$,

$$\begin{cases} z_i = \lambda_i \\ x_{ij} = \lambda_i \beta_j \\ y_{ij} = \lambda_i \beta_j^{-1}. \end{cases}$$

Conversely, starting from a point $(w_1, \dots, w_{2g})$ in $U_{mn}^{\sigma}(F)$ as in (1), the invariance of the statement under permutations allows us to assume that $\sigma = \mathrm{id}$, and we get a decomposition as in (2) by setting $\lambda_i = z_i$ for $i = 1, \dots, 2m$ and $\beta_j = x_{1j}/z_1$ for $j = 1, \dots, \frac{n-1}{2}$. □

Proposition 4.7.5. *Let v be a place of good reduction of A and m, n be integers such that $mn = g$ (with $n \geq 3$). Let $(\mu_1, \dots, \mu_{2g})$ be the eigenvalues of Fr_v and suppose that*

$$(\mu_1, \dots, \mu_{2g}) \notin \bigcup_{\sigma \in S_{2g}} U_{mn}^{\sigma}(\overline{\mathbb{Q}}).$$

Then for every ℓ that is strictly larger than $(2q_v)^{[F(v):\mathbb{Q}]}$ the element $\rho_\ell(\mathrm{Fr}_v)$ does not lie in a tensor product subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ of type (m, n). In particular, for any such ℓ the group G_ℓ is not contained in a tensor product subgroup of type (m, n).

Proof. Since clearly $\prod_{i=1}^{2g} \mu_i \neq 0$, the fact that $(\mu_1, \dots, \mu_{2g})$ does not belong to $U_{mn}^{\sigma}\left(\overline{\mathbb{Q}}\right)$ for any σ is equivalent to the fact that for every $\sigma \in S_{2g}$ there is a $p^{\sigma} \in \mathcal{P}_{mn}^{\sigma}$ (cf. definition 4.7.3) such that $\alpha_p^{\sigma} := p^{\sigma}(\mu_1, \dots, \mu_{2g})$ is nonzero; recall that p^{σ} is a homogeneous binomial of degree 2 with coefficients in $\{\pm 1\}$. Since the μ_i's are algebraic integers, so are the α_p^{σ}; furthermore, every α_p^{σ} belongs to $F(v)$, the splitting field of $f_v(x)$. Finally, the absolute value of every Galois conjugate of every μ_i is $q_v^{1/2}$, so $|\alpha_p^{\sigma}| \leq 2q_v$ under any embedding of $F(v)$ in $\mathbb{C}$: putting everything together we see that, for every fixed σ, the set of numbers $\left\{ a_p^{\sigma} := N_{F(v)/\mathbb{Q}}\left(\alpha_p^{\sigma}\right) \;\middle|\; p \in \mathcal{P}_{mn}^{\sigma} \right\}$ consists of integers of absolute value at most $(2q_v)^{[F(v):\mathbb{Q}]}$, not all equal to zero. Suppose now by contradiction that

$\rho_\ell(\mathrm{Fr}_v)$ lies in a tensor product subgroup of type (m,n). By lemma 4.7.2, the eigenvalues $\overline{\mu_1},\dots,\overline{\mu_{2g}}$ of $\rho_\ell(\mathrm{Fr}_v)$ can be written as

$$\left\{\overline{\lambda_i},\ \overline{\lambda_i}\cdot\overline{\beta_j},\ \overline{\lambda_i}\cdot\overline{\beta_j}^{-1}\ \middle|\ i=1,\dots,2m,\ j=1,\dots,\frac{n-1}{2}\right\}$$

for some elements $\overline{\lambda_i},\overline{\beta_j}$ of $\overline{\mathbb{F}_\ell}^\times$, and by lemma 4.7.4 there is a permutation σ such that $(\overline{\mu_1},\dots,\overline{\mu_{2g}})$ defines a point of $U^\sigma_{mn}(\overline{\mathbb{F}_\ell})$. This implies that (for this specific choice of σ) all the numbers a_p^σ reduce to 0 in $\overline{\mathbb{F}_\ell}$, and since the a_p^σ are integers this amounts to saying that ℓ divides all the a_p^σ (for $p\in\mathcal{P}^\sigma_{mn}$). However, we have seen that there is at least one polynomial $p\in\mathcal{P}^\sigma_{mn}$ for which a_p^σ is nonzero, so $\ell\mid a_p^\sigma$ implies $\ell\le|a_p^\sigma|\le(2q_v)^{[F(v):\mathbb{Q}]}$: this clearly contradicts our choice of ℓ, and the proposition is proved. □

Serre has proved [120, p. 49] that places v as in the statement of the proposition do exist, and in fact a slight modification of his argument shows that they have density 1. On the other hand, the following lemma gives an easily testable (sufficient) criterion to decide whether or not a place v satisfies the hypotheses of the previous proposition:

Lemma 4.7.6. *Let v be a place of K of good reduction for A such that the Galois group of $f_v(x)$ is the full Weyl group $\mathcal{W}_g:=(\mathbb{Z}/2\mathbb{Z})^g\rtimes S_g$. Let $(\mu_1,\dots,\mu_{2g})$ be the eigenvalues of Fr_v. Then for any choice of positive integers (m,n) with $n\ge3$ and $mn=g$ the point $(\mu_1,\dots,\mu_{2g})$ does not belong to $\bigcup_{\sigma\in S_{2g}}U^\sigma_{mn}(\overline{\mathbb{Q}})$.*

Proof. Let $s_v(x)\in\mathbb{Z}[x]$ be the squarefree part of $f_v(x)$ and s be its degree. Like $f_v(x)$, the polynomial $s_v(x)$ satisfies $x^s s_v(q/x)=q^{s/2}s_v(x)$, so its Galois group is a subgroup of $(\mathbb{Z}/2\mathbb{Z})^s\rtimes S_s$: as the splitting fields of $s_v(x)$ and $f_v(x)$ coincide, we must have $s=g$, that is, the μ_i's are all distinct.

Let now λ,ν_1,ν_2 be any three distinct eigenvalues of Fr_v. We shall show that we cannot have $\lambda^2=\nu_1\nu_2$, hence in particular no permutation of the μ_i's can define a point of $U_{mn}(\overline{\mathbb{Q}})$ (recall that one of the equations defining U_{mn} is $z_1^2=x_{11}y_{11}$). Suppose by contradiction that $\lambda^2=\nu_1\nu_2$. Up to renumbering the μ_i's, the action of $\mathcal{W}_g$ on the set $\{\mu_1,\dots,\mu_{2g}\}$ has the following property: for every $\sigma\in\mathcal{W}_g$ and for every pair of indices i,j, we have $\sigma(\mu_i)=\mu_j$ if and only if $\sigma(\mu_{2g+1-i})=\mu_{2g+1-j}$. We call μ_{2g+1-i} the conjugate of μ_i. Suppose first that ν_2 is not the conjugate of ν_1, nor of λ: then there exists a $\sigma\in\mathcal{W}_g$ which fixes both ν_1 and λ, but such that $\sigma(\nu_2)\ne\nu_2$. Applying σ to the equality $\lambda^2=\nu_1\nu_2$ we find $\lambda^2=\nu_1\sigma(\nu_2)$, which is a contradiction since $\sigma(\nu_2)\ne\nu_2$. Next suppose that ν_2 is the conjugate of λ: then ν_1 is *not* the conjugate of λ, nor of ν_2 (since λ,ν_1,ν_2 are all distinct), and we can just repeat the same argument with ν_2 replaced by ν_1. Finally, assume ν_1,ν_2 are conjugate to each other (hence not to λ), and denote by S the stabilizer of ν_1,ν_2 in $\mathcal{W}_g$: since $g\ge3$, the orbit of λ under the action of S has order at least 4, hence in particular there is a $\sigma\in S$ such that $\sigma(\lambda)\ne\pm\lambda$. Applying this σ to the equation $\lambda^2=\nu_1\nu_2$ leads once more to a contradiction. □

4.8 Proof of theorem 4.1.3

It is clear that the prime ℓ is larger than $g(2g-1)+1$ (cf. proposition 4.5.9), than $(2g+1)(2g+2)$ (cf. proposition 4.6.4), than $J(2g)+1$ (cf. corollary 4.6.6) and than $\frac{1}{2}(2g+1)^{12g}$ (the bound appearing

in property (∗)). By corollary 4.2.4 we see that it is enough to show that G_ℓ contains $\mathrm{Sp}_{2g}(\mathbb{F}_\ell)$, so suppose this is not the case: G_ℓ is then contained in one of the maximal subgroups of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ listed in theorem 4.3.14. Let us go through this list. Given the inequalities imposed on ℓ, proposition 4.4.1 implies that cases 1 through 3 cannot happen. Likewise proposition 4.7.5 (which can be applied thanks to lemma 4.7.6) implies that case 4 cannot arise, so we are left with considering the case of G_ℓ being contained in a maximal subgroup G of class $\mathcal{S}$. If $\operatorname{soc}\mathbb{P}G$ is of Lie type in characteristic ℓ, then property (∗) implies that $\operatorname{soc}\mathbb{P}G \cong \mathrm{PSL}_2(\mathbb{F}_\ell)$, which by proposition 4.5.9 cannot happen for $\ell > g(2g-1)+1$. If, on the contrary, $\operatorname{soc}\mathbb{P}G$ is not of Lie type in characteristic ℓ, then proposition 4.6.4 implies that $\mathbb{P}G$ is of order at most $J(2g)$, which is impossible by corollary 4.6.6. Finally, it is clear from the explicit expressions of $b([K:\mathbb{Q}], g, h(A))$ that the function $b(g!\cdot[K:\mathbb{Q}], g, h(A))$ grows faster than $b([K:\mathbb{Q}], 2g, 2h(A))^{1/2g}$, and it is easy to check that for $g \geq 19$ the inequality $b(A/K; g!) > b(A^2/K; g)^{1/2g}$ holds for any K and any A. □

Remark 4.8.1. Notice that all that is used about v is that Fr_v satisfies the hypothesis of proposition 4.7.5 for all pairs (m,n) such that $mn = g$; we shall need this fact in the next section.

4.9 The tensor product case II

In this section we show that, when $\dim(A) = 3$, a place v satisfying the hypothesis of proposition 4.7.5 can be found whose residue characteristic is bounded explicitly in terms of simple arithmetical invariants of A/K. This will be achieved through an application of Chebotarev's theorem, but we shall first need a certain number of preliminaries. We continue using the notation of §4.7; in particular, if v is a finite place of K we denote by p_v (resp. q_v) the characteristic (resp. the cardinality) of the residue field at v. We also introduce the set

$$\Omega_K^A := \left\{ v \in \Omega_K \;\middle|\; A \text{ has good reduction at } v \text{ and } v \text{ has degree 1 over } \mathbb{Q} \right\}.$$

Most of what we do in this section could be generalized to some extent to other values of g: for example, all results up to corollary 4.9.11 can easily be extended to cover the case of an arbitrary (odd) *prime* dimension, and it is only the proof of proposition 4.9.12 that depends on the assumption $\dim A = 3$, since it relies on the particularly simple subgroup structure of $\mathrm{CGO}_3(\mathbb{F}_\ell)$. Trying to generalize this result to other $g \geq 5$, one is faced with problems akin to those that forced us to impose condition (∗) on the dimension g: the group $\mathrm{GL}_2(\mathbb{F}_\ell) \otimes \mathrm{CGO}_g(\mathbb{F}_\ell)$ contains families of maximal proper subgroups of Lie type which we cannot exclude by simply looking at the action of inertia on $A[\ell]$.

Comparing our arguments with those used by Serre [121] to prove his open image theorem for abelian varieties of odd dimension with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, it is easy to realize that a major stumbling block in our approach is the fact that there is no clear analogue of Sen's theorem [114] for representations over $\mathbb{F}_\ell$: indeed, Sen's theorem strongly depends on the completeness of $\mathbb{C}_p$, and it is not even clear what a modulo-ℓ analogue of this theorem should look like.

4.9.1 Decompositions of the eigenvalues of Fr_v

We start with two easy lemmas, which do not depend on the assumption $\dim A = 3$:

Lemma 4.9.1. *Let N be a positive integer no less than 3. Suppose all the torsion points of A of order N are defined over K, and let v be any place of K of good reduction for A and not dividing N. The group generated by the eigenvalues of Fr_v does not contain any nontrivial root of unity.*

Proof. Let $\mu_1, \dots, \mu_{2g}$ be the eigenvalues of Fr_v. Looking at the action of Fr_v on $A[N]$ we see that each of them (hence every element of the group they generate) is congruent to 1 modulo N, but as it is well known there are no nontrivial roots of unity congruent to 1 modulo N when $N \geq 3$. □

Lemma 4.9.2. *Let N be a positive integer no less than $2g+1$. Suppose all the torsion points of A of order N are defined over K, and let v be a place in Ω_K^A. If p_v does not divide N and is larger than $(2g)^2$, then p_v does not divide $\operatorname{tr} \mathrm{Fr}_v$.*

Proof. On the one hand $\mathrm{Gal}(\overline{K}/K)$ acts trivially on $A[N]$, so $\operatorname{tr} \mathrm{Fr}_v$ cannot be zero since it is congruent to $2g$ modulo N. On the other hand, the Weil conjectures imply that $|\operatorname{tr} \mathrm{Fr}_v|$ does not exceed $2g \cdot p_v^{1/2}$, so if p_v divides $|\operatorname{tr} \mathrm{Fr}_v| \neq 0$ we must have $p_v \leq 2g \cdot p_v^{1/2}$, which is equivalent to $p_v \leq (2g)^2$. □

We now specialize to the case $\dim A = 3$. Notice that all tensor product subgroups of $\mathrm{GSp}_6(\mathbb{F}_\ell)$ are of type $(1,3)$, that is, up to conjugation they can be identified with the group $\mathrm{GL}_2(\mathbb{F}_\ell) \otimes \mathrm{CGO}_3(\mathbb{F}_\ell)$. The following proposition imposes stringent restrictions on a Frobenius whose eigenvalues define a point of $U_{13}(\overline{\mathbb{Q}})$:

Proposition 4.9.3. *Let N be an integer no less than $2g+1 = 7$. Suppose all the torsion points of A of order N are defined over K and let v be a place of K that satisfies:*

- *$v \in \Omega_K^A$ and $p_v > \max\{N, (2g)^2\}$;*
- *the eigenvalues $(\mu_1, \dots, \mu_{2g})$ of Fr_v define a point of $\bigcup_{\sigma \in S_{2g}} U_{13}^\sigma(\overline{\mathbb{Q}})$, i.e. Fr_v does not satisfy the hypothesis of proposition 4.7.5.*

Then at least one of the following holds:

1. *there exist algebraic integers λ_1, λ_2 such that the eigenvalues of Fr_v are given by λ_1 and λ_2, both with multiplicity $g = 3$;*
2. *for any choice of elements $\lambda_1, \lambda_2, \beta$ of $\overline{\mathbb{Q}}^\times$ such that the multisets $\{\lambda_i\beta, \lambda_i, \lambda_i\beta^{-1} \mid i = 1,2\}$ and $\{\mu_1, \dots, \mu_{2g}\}$ coincide, the algebraic number $\lambda_1 + \lambda_2$ is not an integer (at least one valid choice of λ_i, β exists by lemma 4.7.4).*

Proof. Notice first that, by lemma 4.9.2, the residue characteristic p_v does not divide the (nonzero) integer $\operatorname{tr} \mathrm{Fr}_v$. Let now λ_1, λ_2 and β be algebraic numbers such that the eigenvalues of Fr_v are λ_1, λ_2 and $\lambda_i\beta^{\pm 1}$ for $i = 1, 2$. As the eigenvalues of Fr_v are algebraic integers, this implies in particular that λ_1, λ_2 are algebraic integers. If $\lambda_1 + \lambda_2$ is not an integer for any choice of λ_i, β we are done, hence (without loss of generality) we can work under the additional assumption that $\lambda_1 + \lambda_2$ is an integer. We are thus reduced to showing that $\beta = 1$: this we shall do by proving that β is a root of unity, and then applying lemma 4.9.1. Let w be any place of $\overline{\mathbb{Q}}$. Suppose first that the residual

characteristic of w is not p_v: the Weil conjectures imply that the eigenvalues of Fr_v are units away from p_v, hence $\mathrm{ord}_w(\lambda_i\beta) = \mathrm{ord}_w(\lambda_i\beta^{-1}) = 0$, which immediately gives $\mathrm{ord}_w(\beta) = 0$.
Suppose now that the residual characteristic of w is p_v. As $\operatorname{tr}\mathrm{Fr}_v \neq 0$ can also be written as $(\lambda_1+\lambda_2)\left(1+\beta+\beta^{-1}\right)$ we see that $\lambda_1+\lambda_2$ is nonzero. If $\mathrm{ord}_w(\lambda_i)$ is positive for $i = 1, 2$, then $\mathrm{ord}_w\left(\sum_i \lambda_i\right)$ is positive as well and therefore (since $\lambda_1+\lambda_2$ is an integer) we see that p_v divides $\lambda_1+\lambda_2$. However, the Weil conjectures also imply that $|\lambda_1+\lambda_2| \leq 2\sqrt{p_v}$, which – combined with the fact that $\lambda_1+\lambda_2$ is nonzero – gives a contradiction for $p_v \geq 5$ (and our assumptions entail in particular $p_v > (2g)^2 = 36$), so without loss of generality we can assume $\mathrm{ord}_w(\lambda_1) = 0$. Now since $\lambda_1\beta$ and $\lambda_1\beta^{-1}$ are algebraic integers they both have non-negative valuation at w, so we have

$$0 \leq \mathrm{ord}_w(\lambda_1\beta) = \mathrm{ord}_w(\beta), \quad 0 \leq \mathrm{ord}_w(\lambda_1\beta^{-1}) = -\mathrm{ord}_w(\beta),$$

and therefore $\mathrm{ord}_w(\beta) = 0$. It follows that the algebraic number β has zero valuation at all places of $\overline{\mathbb{Q}}$ and is therefore a root of unity; by lemma 4.9.1, this implies $\beta = 1$. □

We now proceed to give a sufficient criterion for case (2) of the previous proposition not to happen. The criterion is not new, and can be deduced for example from [21, Sublemmas 5.2.3 and 5.2.4]; however, given that our setting is slightly different and the statement itself differs from Chi's, we reproduce the argument in full for the reader's convenience. Before discussing the criterion itself we set up some notation.

Definition 4.9.4. We say that a Frobenius element Fr_v is **of tensor product type** if the multiset Δ of eigenvalues of Fr_v can be written as

$$\Delta = \left\{\lambda_i, \lambda_i\beta^{\pm 1} \;\middle|\; i = 1, 2\right\}$$

for some choice of λ_i, β in $\overline{\mathbb{Q}}^\times$. When this is the case, we write Ψ (resp. Λ) for the multiset $\left\{1, \beta^{\pm 1}\right\}$ (resp. $\{\lambda_1, \lambda_2\}$), and we also write symbolically $\Delta = \Lambda \cdot \Psi$.

Remark 4.9.5. *A priori*, the eigenvalues of Fr_v could admit more than one decomposition as in the previous definition. We shall be careful to distinguish those statements that hold for *any* such decomposition from those that hold for a *fixed* decomposition. Also notice that lemma 4.7.4 amounts to saying that a Frobenius Fr_v is of tensor product type if and only if its eigenvalues define a point of $\bigcup_{\sigma\in S_{2g}} U_{13}^{\sigma}(\overline{\mathbb{Q}})$.

We now introduce a weak notion of multiplicative independence for the eigenvalues of a Frobenius Fr_v of tensor product type. Fix sets Λ and Ψ as in definition 4.9.4, and consider the equation

$$(x_1\psi_1)^2 = (x_2\psi_2)(x_3\psi_3) \tag{4.5}$$

in unknowns $x_1, x_2, x_3 \in \Lambda$ and $\psi_1, \psi_2, \psi_3 \in \Psi$. Notice that this equation admits two obvious families of solutions: if we take $x_1 = x_2 = x_3$, the equation reduces to $\psi_1^2 = \psi_2\psi_3$, which for all $\psi \in \Psi$ admits the solutions $1^2 = \psi\cdot\psi^{-1}$ and $\psi^2 = \psi\cdot\psi$; if no other solution exists, we say that the eigenvalues of Fr_v are weakly independent. More precisely, we give the following definition:

Definition 4.9.6. We say that the eigenvalues of Fr_v are **weakly independent** (with respect to a given decomposition of $\Delta = \Lambda\cdot\Psi$) if the following two conditions hold:

1. the eigenvalues of Fr_v are all distinct;

2. if $(x_1, x_2, x_3, \psi_1, \psi_2, \psi_3) \in \Lambda^3 \times \Psi^3$ is a solution to equation (4.5), then $x_1 = x_2 = x_3$ and there exists $\psi \in \Psi$ such that either $(\psi_1, \psi_2, \psi_3) = (1, \psi, \psi^{-1})$ or $(\psi_1, \psi_2, \psi_3) = (\psi, \psi, \psi)$.

A first useful feature of the notion of weak independence is that it entails unicity of the decomposition $\Delta = \Lambda \cdot \Psi$:

Lemma 4.9.7. *Suppose that* Fr_v *is of tensor product type and that its eigenvalues are weakly independent with respect to a certain decomposition* $\Delta = \Lambda \cdot \Psi$*: then* $\lambda_1 + \lambda_2$ *is an integer, and for any decomposition* $\Delta = \Lambda' \cdot \Psi'$ *of* Δ *we have* $\Lambda' = \Lambda$ *and* $\Psi = \Psi'$*.*

Proof. We start by describing a property that characterizes λ_1, λ_2 among the elements of Δ. For every $\gamma \in \Delta$ we consider the map

$$\begin{array}{rccc} T_\gamma : & \Delta & \to & \overline{\mathbb{Q}}^\times \\ & \delta & \mapsto & \frac{\gamma^2}{\delta}. \end{array}$$

Claim. We have $|T_\gamma(\Delta) \cap \Delta| \geq g = 3$ if and only if γ belongs to Λ.

Proof of claim. The "if" part is trivial: if $\gamma = \lambda_i$, then it is clear that $T_{\lambda_i}(\lambda_i \psi) \in \Delta$ for all $\psi \in \Psi$; as T_γ is injective, this gives $|\Psi| = 3$ elements in the intersection $T_\gamma(\Delta) \cap \Delta$.
Conversely, suppose that $|T_\gamma(\Delta) \cap \Delta| \geq 3$ for a certain $\gamma \in \Delta$. Write $\gamma = x_1\psi_1$ with $x_1 \in \Lambda, \psi_1 \in \Psi$ and suppose $\psi_1 \neq 1$. Let $x_2\psi_2 \in \Delta$ be such that $T_\gamma(x_2\psi_2) \in \Delta$. By definition, this implies the existence of $x_3 \in \Lambda, \psi_3 \in \Psi$ that satisfy

$$\frac{(x_1\psi_1)^2}{x_2\psi_2} = x_3\psi_3,$$

and since the eigenvalues are weakly independent we have $x_2 = x_1$ and $\psi_2 = \psi_1$ (since $\psi_1 \neq 1$). Hence we see that $\lambda_1\psi_1$ is the only eigenvalue δ of Fr_v such that $T_\gamma(\delta)$ belongs to Δ, contradicting the fact that $|T_\gamma(\Delta) \cap \Delta| \geq g = 3$.
Notice now that λ_1 and λ_2, being eigenvalues of Fr_v, are algebraic integers, so in order to show that $\lambda_1 + \lambda_2$ is an integer it suffices to prove that it is a rational number, i.e. that the set $\{\lambda_1, \lambda_2\}$ is $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-invariant. By the previous characterization of λ_1, λ_2 it then suffices to show that for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ we have $\left|T_{\sigma(\lambda_i)}(\Delta) \cap \Delta\right| \geq g = 3$, and this follows from

$$\left|T_{\sigma(\lambda_i)}(\Delta) \cap \Delta\right| = \left|T_{\sigma(\lambda_i)}(\sigma(\Delta)) \cap \sigma(\Delta)\right| = |T_{\lambda_i}(\Delta) \cap \Delta| \geq g = 3,$$

where we have used the equality $\sigma(\Delta) = \Delta$ (the set Δ is $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-stable since the characteristic polynomial of Fr_v has integral coefficients).
Moreover, the characterization we have given of λ_1, λ_2 does not use the decomposition of Δ we have fixed, hence it uniquely determines the values of λ_1, λ_2 in any possible decomposition $\Delta = \Lambda' \cdot \Psi'$.
We show that the set Ψ is uniquely determined as well. Let $\Delta = \Lambda \cdot \Psi'$ be any decomposition of Δ, with $\Psi' = \left\{1, (\beta')^{\pm 1}\right\}$, and suppose that $\beta' \neq \beta^{\pm 1}$. By definition, $\mu = \lambda_1\beta$ is an element of Δ, hence it can be written as $\mu = \lambda_i\psi'$ for some $\psi' \in \Psi'$ and some $i \in \{1, 2\}$. As the eigenvalues of Fr_v are all distinct we necessarily have $\psi' \neq 1$; furthermore, if we had $i = 1$ we would also have $\psi' = \beta$, a contradiction, so (replacing β' by $(\beta')^{-1}$ if necessary) we must in fact have $\mu = \lambda_2\beta'$. It follows that β' is equal to $\frac{\lambda_1}{\lambda_2}\beta$ and hence Δ also contains $\lambda_1\beta' = \frac{\lambda_1^2}{\lambda_2}\beta$, which in turn must be of the form $\lambda_k\psi$ for some $k \in \{1, 2\}$ and $\psi \in \Psi$. Thus we find that $\frac{(\lambda_1\beta)^2}{\lambda_2\beta} = \lambda_k\psi$ is a solution to equation (4.5), so by definition of weak independence we must have $\lambda_1 = \lambda_2$, which is absurd since the eigenvalues of Fr_v are all distinct. The contradiction shows that $\beta' = \beta$, that is, $\Psi' = \Psi$. □

We also need a version of definition 4.9.6 for operators acting on $\mathbb{F}_\ell^{2g}$:

Definition 4.9.8. Let h be an element of $\mathrm{GL}_{2g}(\mathbb{F}_\ell)$. If the multiset Δ_ℓ of eigenvalues of h in $\overline{\mathbb{F}_\ell}^\times$ can be written as $\Lambda_\ell \cdot \Psi_\ell$, where $\Lambda_\ell = \{\lambda_1, \lambda_2\}$ and $\Psi_\ell = \{1, \beta^{\pm 1}\}$ for some $\lambda_i, \beta \in \overline{\mathbb{F}_\ell}^\times$, we say that h is of tensor product type (modulo ℓ). If furthermore the elements of Δ_ℓ are all distinct, and the equality $(x_1\psi_1)^2 = (x_2\psi_2)(x_3\psi_3)$ with $x_i \in \Lambda_\ell, \psi_j \in \Psi_\ell$ implies $x_1 = x_2 = x_3$ and either $(\psi_1, \psi_2, \psi_3) = (1, \psi, \psi^{-1})$ or $(\psi_1, \psi_2, \psi_3) = (\psi, \psi, \psi)$ for some $\psi \in \Psi_\ell$, then we say that h has weakly independent eigenvalues modulo ℓ.

As the proof of lemma 4.9.7 does not use any particular features of the field $\mathbb{Q}$, the same argument also shows:

Lemma 4.9.9. *Suppose $h \in \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ is of tensor product type and has weakly independent eigenvalues modulo ℓ: then the decomposition $\Delta_\ell = \Psi_\ell \cdot \Lambda_\ell$ is unique.*

Lemma 4.9.10. *Let v be a place in Ω_K^A. Suppose that Fr_v is of tensor product type and ℓ is a prime different from p_v: then $\rho_\ell(\mathrm{Fr}_v)$ is of tensor product type. If furthermore $\rho_\ell(\mathrm{Fr}_v)$ has weakly independent eigenvalues modulo ℓ (for some, hence for any, decomposition of Δ_ℓ as $\Lambda_\ell \cdot \Psi_\ell$), then Fr_v has weakly independent eigenvalues as well. In particular, the decomposition $\Delta = \Lambda \cdot \Psi$ of the eigenvalues of Fr_v is unique, and it satisfies $\lambda_1 + \lambda_2 \in \mathbb{Z}$.*

Proof. The first statement is clear: a decomposition of the eigevanlues of Fr_v induces an analogous decomposition of the eigenvalues of $\rho_\ell(\mathrm{Fr}_v)$. As for the second part, notice first that by assumption the eigenvalues of $\rho_\ell(\mathrm{Fr}_v)$ are distinct, hence the eigenvalues of Fr_v are a fortiori distinct, and there is a *unique* way to lift an eigenvalue of $\rho_\ell(\mathrm{Fr}_v)$ to an eigenvalue of Fr_v. Denote by Δ (resp. Δ_ℓ) the set of eigenvalues of Fr_v (resp. of $\rho_\ell(\mathrm{Fr}_v)$); by assumption, there exists a decomposition $\Delta = \Lambda \cdot \Psi$, which induces an analogous decomposition $\Delta_\ell = \Lambda_\ell \cdot \Psi_\ell$. The multiset Δ does not contain elements with multiplicity greater than 1, so the map

$$\begin{array}{ccc} \Lambda \times \Psi & \to & \Delta \\ (\lambda, \psi) & \mapsto & \lambda\psi \end{array}$$

is a bijection: equivalently, for every eigenvalue δ of Fr_v, *in the given decomposition* $\Lambda \cdot \Psi$ there exist unique $\lambda \in \Lambda$ and $\psi \in \Psi$ such that $\delta = \lambda \cdot \psi$. Repeating the same argument modulo ℓ we find that $\Psi \times \Lambda \to \Delta \to \Delta_\ell \to \Psi_\ell \times \Lambda_\ell$ is a bijection. Consider now the equation

$$(x_1\psi_1)^2 = (x_2\psi_2)(x_3\psi_3)$$

with $x_i \in \Lambda$ and $\psi_j \in \Psi$. Reducing modulo ℓ and using the weak independence of the eigenvalues of $\rho_\ell(\mathrm{Fr}_v)$ we see that $x_1 = x_2 = x_3$ (as elements of Λ_ℓ), and either $\psi_1 = \psi_2 = \psi_3$ or $\psi_1 = 1$ and $\psi_2 = \psi_3^{-1}$ (as elements of Ψ_ℓ). Using the fact that $\Psi \times \Lambda \to \Psi_\ell \times \Lambda_\ell$ is a bijection we then conclude that we also have $x_1 = x_2 = x_3$ as elements of Λ, and that (ψ_1, ψ_2, ψ_3) is either of the form $(1, \psi, \psi^{-1})$ or of the form (ψ, ψ, ψ) for some $\psi \in \Psi$. The remaining statements follow immediately from lemma 4.9.7. □

We finally come to the result which will allow us to find Frobenius elements not of tensor product type:

Corollary 4.9.11. *Let N be an integer no less than $2g+1=7$. Suppose that all the torsion points of A of order N are defined over K, and let $v \in \Omega_K^A$ satisfy $p_v > \max\{N, (2g)^2\}$. Suppose furthermore that for some prime ℓ different from p_v the image $\rho_\ell(\mathrm{Fr}_v)$ is of tensor product type and has weakly independent eigenvalues modulo ℓ. Then Fr_v is **not** of tensor product type.*

Proof. Suppose Fr_v is of tensor product type: then it satisfies the assumptions of lemma 4.9.10, so in the (unique) decomposition of its eigenvalues as $\Lambda \cdot \Psi$ we must have $\lambda_1 + \lambda_2 \in \mathbb{Z}$. Furthermore, the eigenvalues of Fr_v are all distinct (since this is true when they are regarded modulo ℓ). On the other hand, Fr_v also satisfies the hypotheses of proposition 4.9.3, hence one of the two conclusions of that proposition must hold: but this is absurd by what we just proved, and the contradiction shows the result. □

We now just need to find a Frobenius Fr_v as in the previous corollary: this will be achieved by an application of Chebotarev's theorem, for which we need one more lower bound on G_ℓ (recall that the group $\Omega_3(\mathbb{F}_\ell)$ was introduced in definition 4.3.9):

Proposition 4.9.12. *Suppose that the 7-torsion of A is defined over K: then for all primes ℓ unramified in K and strictly larger than $b(A^2/K;3)^{1/6}$ we have $G_\ell \supseteq \mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$.*

Proof. This is very similar to what we did in the previous sections, so we keep details to a minimum. Notice first that we can assume that (up to conjugation) G_ℓ is contained in $\mathrm{GL}_2(\mathbb{F}_\ell) \otimes \mathrm{CGO}_3(\mathbb{F}_\ell)$, for otherwise the proof of theorem 4.1.3 shows that G_ℓ contains all of $\mathrm{Sp}_6(\mathbb{F}_\ell)$. Also notice that the group $\mathrm{GL}_2(\mathbb{F}_\ell) \otimes \mathrm{CGO}_3(\mathbb{F}_\ell)$ admits well-defined projections π_2, π_3 to $\mathrm{PGL}_2(\mathbb{F}_\ell)$ and $\mathbb{P}\mathrm{CGO}_3(\mathbb{F}_\ell)$ respectively. Also notice that the tensor product structure implies that if either projection stabilizes a subspace (respectively in $\mathbb{F}_\ell^2$ or in $\mathbb{F}_\ell^3$), then the same is true for all of G_ℓ: indeed, if W is a point of $\mathbb{P}\left(\mathbb{F}_\ell^2\right)$ (i.e. a line in $\mathbb{F}_\ell^2$) stable under the action of $\pi_2(G_\ell)$, then $W \otimes \mathbb{F}_\ell^3$ is a proper subspace of $\mathbb{F}_\ell^6$ stable under the action of G_ℓ, and the same argument applies to π_3 as well. In particular, proposition 4.4.1 implies that neither projection stabilizes a linear subspace. We now show that the two projections are in fact surjective.

Surjectivity on $\mathbb{P}\Omega_3(\mathbb{F}_\ell) \cong \mathrm{PSL}_2(\mathbb{F}_\ell)$. From [19, Table 8.7] we see that the maximal subgroups of $\mathbb{P}\mathrm{CGO}_3(\mathbb{F}_\ell)$ that do not contain $\mathbb{P}\Omega_3$ either stabilize a linear subspace or have order at most 120. We have already excluded the first case, and the second case is easily treated as well: replacing K with the extension defined by $\ker\left(\mathrm{Gal}(\overline{K}/K) \to G_\ell \to \mathbb{P}\mathrm{CGO}_3(\mathbb{F}_\ell)\right)$ we are back to the case of a group stabilizing a linear subspace, hence this case cannot happen for ℓ in our range (since we have in particular $\ell > b_0(A/K;120)$).

Remark 4.9.13. Notice that although $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$ and $\mathrm{PSL}_2(\mathbb{F}_\ell)$ are isomorphic as abstract groups, the representation structure of their respective natural modules is very different: in particular, the non-split Cartan subgroups are of class $\mathcal{C}_3$ in $\mathrm{PSL}_2(\mathbb{F}_\ell)$ but of class $\mathcal{C}_1$ in $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$.

(Almost) surjectivity on $\mathrm{PSL}_2(\mathbb{F}_\ell)$. We read from [19, Table 8.1] that the maximal subgroups of $\mathrm{PGL}_2(\mathbb{F}_\ell)$ that do not contain $\mathrm{PSL}_2(\mathbb{F}_\ell)$ and do not stabilize a linear subspace either contain a normal abelian subgroup of index at most 2, or have order at most 120. The second case is excluded by the same argument as in the previous paragraph, so the image H_2 of G_ℓ in $\mathrm{PGL}_2(\mathbb{F}_\ell)$ contains either $\mathrm{PSL}_2(\mathbb{F}_\ell)$ or an abelian subgroup C_2 of index at most 2; furthermore, in the latter case there

is no loss of generality in assuming that $|C_2| > 60$ (for otherwise H_2 has order at most 120, which is excluded).

Surjectivity on both factors. Let $H_2 = \pi_2(G_\ell), H_3 = \pi_3(G_\ell)$. We consider the image of G_ℓ in $\mathrm{PGL}_2(\mathbb{F}_\ell) \times \mathbb{P}\mathrm{CGO}_3(\mathbb{F}_\ell)$: it is a group $H \subseteq H_2 \times H_3$ that projects surjectively on the factors H_2, H_3. We also know that H_3 contains $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$. Suppose by contradiction that H_2 contains an abelian subgroup C_2 of index at most 2 and replace K with its (at most) quadratic extension K' defined by $\ker\big(\mathrm{Gal}(\overline{K}/K) \to G_\ell \to H_2 \to H_2/C_2\big)$. This has the effect of replacing H_2 with C_2; at the same time H_3 gets replaced by a subgroup C_3 of index at most 2, and since $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$ does not have subgroups of index 2 we see that $C_3 \supseteq \mathbb{P}\Omega_3(\mathbb{F}_\ell)$. Finally, G_ℓ is replaced by a subgroup $\tilde{G}_\ell$ of index at most 2, and likewise H gets replaced by a subgroup C of index at most 2, which satisfies $C \subseteq C_2 \times C_3$ and projects surjectively on both C_2 and C_3. Let now $N_3 := \ker(C \to C_2)$ and $N_2 := \ker(C \to C_3)$, considered as subgroups of C_3, C_2 respectively. By Goursat's lemma we know that the quotients C_3/N_3 and C_2/N_2 are isomorphic, and in particular abelian (as C_2 is). Since the group $\mathbb{P}\mathrm{CGO}_3(\mathbb{F}_\ell)$ is almost simple with socle $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$, it is clear that N_3 contains all of $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$, so the quotient C_3/N_3 has order at most 2. Hence N_2 has in turn index at most 2 in C_2, and therefore there is a nontrivial element α in N_2 (recall that $|C_2| > 60$). By definition of N_2, this α projects to the identity in C_3, so any element $\tilde{\alpha} \in \tilde{G}_\ell$ lifting α is central in $\tilde{G}_\ell$. In particular, the centralizer of $\tilde{G}_\ell$ in $\mathrm{Aut}\, A[\ell]$ is larger than $\mathbb{F}_\ell$, and by lemma 3.3.17 this is a contradiction for ℓ larger than $b(A^2/K')^{1/6}$, a quantity which is smaller than $b(A^2/K;3)^{1/6}$.

G_ℓ contains $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$. Notice that it is enough to show that H (the image of $\pi_2 \times \pi_3$) contains $\mathrm{PSL}_2(\mathbb{F}_\ell) \times \mathbb{P}\Omega_3(\mathbb{F}_\ell)$. Indeed, if this is the case, then for every $\overline{x_2} \in \mathrm{PSL}_2(\mathbb{F}_\ell)$ we can find an $x \in G_\ell$ with $\pi_2(x) = \overline{x_2}$ and $\pi_3(x) = \mathrm{Id}$, that is G_ℓ contains a certain x that can be written as $x = x_2 \otimes \mathrm{Id}$ for some $x_2 \in \mathrm{GL}_2(\mathbb{F}_\ell)$ lifting $\overline{x_2}$. Consider now the subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$ given by $\big\{x \in \mathrm{GL}_2(\mathbb{F}_\ell) \;\big|\; x \otimes \mathrm{Id} \in G_\ell\big\}$: by what we just said, this group projects surjectively onto $\mathrm{PSL}_2(\mathbb{F}_\ell)$, hence it contains all of $\mathrm{SL}_2(\mathbb{F}_\ell)$. It follows that G_ℓ contains $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \{\mathrm{Id}\}$, and by the same argument applied to π_3 we also have $\{\mathrm{Id}\} \otimes \Omega_3(\mathbb{F}_\ell) \subseteq G_\ell$, which implies $G_\ell \supseteq \mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$ as claimed.

So let again $H_2 = \pi_2(G_\ell)$ and $H_3 = \pi_3(G_\ell)$, where we now know that H_2 (resp. H_3) contains $\mathrm{PSL}_2(\mathbb{F}_\ell)$ (resp. $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$). Let N_2, N_3 be the kernels of $H \to H_3$, $H \to H_2$ respectively, considered as subgroups of H_2, H_3, and recall that by Goursat's lemma the image of H in $H_2/N_2 \times H_3/N_3$ is the graph of an isomorphism $H_2/N_2 \xrightarrow{\sim} H_3/N_3$. Now N_2 is a normal subgroup of H_2, so either it contains all of $\mathrm{PSL}_2(\mathbb{F}_\ell)$ or it is trivial: in the former case we have $|H_3/N_3| = |H_2/N_2| \leq 2$, which clearly implies that N_3 contains $\mathbb{P}\Omega_3(\mathbb{F}_\ell)$ and H contains $N_2 \times N_3 \supseteq \mathrm{PSL}_2(\mathbb{F}_\ell) \times \mathbb{P}\Omega_3(\mathbb{F}_\ell)$ as claimed. On the other hand, if N_2 is the trivial group then H is the graph of an isomorphism $H_2 \to H_3$; up to conjugation, such an isomorphism is necessarily the 3-dimensional orthogonal projective representation of either $\mathrm{PGL}_2(\mathbb{F}_\ell)$ or $\mathrm{PSL}_2(\mathbb{F}_\ell)$, according to whether H_2 is $\mathrm{PGL}_2(\mathbb{F}_\ell)$ or $\mathrm{PSL}_2(\mathbb{F}_\ell)$. For simplicity of exposition suppose that $H_2 = \mathrm{PSL}_2(\mathbb{F}_\ell)$; the argument is perfectly analogous if $H_2 = \mathrm{PGL}_2(\mathbb{F}_\ell)$. Let σ_2 be the second symmetric power of the standard representation of $\mathrm{SL}_2(\mathbb{F}_\ell)$ (which is also the unique 3-dimensional orthogonal representation of $\mathrm{SL}_2(\mathbb{F}_\ell)$), and recall that if $x \in \mathrm{SL}_2(\mathbb{F}_\ell)$ has eigenvalues λ_1, λ_2, then $\sigma_2(x)$ has eigenvalues $\lambda_1^2, \lambda_1\lambda_2, \lambda_2^2$. Now since $\sigma_2(-\mathrm{Id})$ is trivial σ_2 fits into a diagram

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{F}_\ell) & \xrightarrow{\sigma_2} & \mathrm{CGO}_3(\mathbb{F}_\ell) \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \mathrm{PSL}_2(\mathbb{F}_\ell) & \xrightarrow[\mathbb{P}\sigma_2]{} & \mathbb{P}\mathrm{CGO}_3(\mathbb{F}_\ell), \end{array}$$

and we have just seen that all $h \in H \subseteq H_2 \times H_3$ can be written as $(\pi(x), \mathbb{P}\sigma_2(\pi(x)))$ for some $x \in \mathrm{SL}_2(\mathbb{F}_\ell)$; furthermore, the commutativity of the diagram gives $h = (\pi(x), \pi(\sigma_2(x)))$. Now let $g_2 \otimes g_3$ be an element of G_ℓ (with $g_2 \in \mathrm{GL}_2(\mathbb{F}_\ell), g_3 \in \mathrm{CGO}_3(\mathbb{F}_\ell)$), mapping in H to a certain $h = (\pi(x), \pi(\sigma_2(x)))$: by definition of H, this implies that there are scalars $\nu_2, \nu_3 \in \mathbb{F}_\ell^\times$ such that $g_2 = \nu_2 x$ and $g_3 = \nu_3 \sigma_2(x)$. If we denote λ_1, λ_2 the eigenvalues of x we thus see that the eigenvalues of $g_2 \otimes g_3$ are given by the pairwise products of $\{\nu_2\lambda_1, \nu_2\lambda_2\}$ and $\{\nu_3\lambda_1^2, \nu_3\lambda_1\lambda_2, \nu_3\lambda_2^2\}$; finally letting $\mu = \nu_2\nu_3$, we have proved that the eigenvalues of any $g_2 \otimes g_3 \in G_\ell$ can be written as

$$\{\mu\lambda_1, \mu\lambda_2\} \cdot \{\lambda_1^2, \lambda_1\lambda_2, \lambda_2^2\} = \{\mu\lambda_1^3, \mu\lambda_1^2\lambda_2, \mu\lambda_1\lambda_2^2, \mu\lambda_1^2\lambda_2, \mu\lambda_1\lambda_2^2, \mu\lambda_2^3\} \tag{4.6}$$

for some $\mu \in \mathbb{F}_\ell^\times$ and $\lambda_1, \lambda_2 \in \mathbb{F}_{\ell^2}^\times$. It is clear that the we arrive at the same conclusion also if $H_2 = \mathrm{PGL}_2(\mathbb{F}_\ell)$. To conclude the proof we just need to show that the decomposition of eigenvalues given by (4.6) leads to a contradiction for ℓ large enough, and this can easily be done by the arguments of section 4.5. We give some detail.

Note first that, since we assume that $A[7]$ is defined over K, a theorem of Raynaud [31, Proposition 4.7] implies that A has semistable reduction at all places of characteristic different from 7. In particular, if we let $\mathfrak{l}$ be any place of K of characteristic ℓ, then A has either good or bad semistable reduction at $\mathfrak{l}$, so we can apply theorem 4.5.2. Let $W_1, \dots, W_k$ be the simple Jordan-Hölder quotients of $A[\ell]$ under the action of $I_\mathfrak{l}$ (or equivalently, of $I_\mathfrak{l}^t$). The argument of lemma 4.5.5 implies that every W_i is of dimension at most 2; let m_0 (resp. m_1, m_2) denote the number of simple Jordan-Holder quotients with trivial action of $I_\mathfrak{l}^t$ (resp. with action given by χ_ℓ, by a fundamental character of level 2). Equation (4.2) and lemma 4.5.6 still hold in our present context, and a slight variant of lemma 4.5.7 shows that $m_0 = 0$ for ℓ unramified in K and larger than 7; thus we want to exclude the case $m_2 = 3$. As in the proof of lemma 4.5.8, one sees that the assumption $m_2 = 3$ implies $\lambda_1 = \pm\lambda_2$; on the other hand, for any given $x \in I_\mathfrak{l}^t$ there is a fundamental character of level 2, call it φ, such that $\mu\lambda_1^3 = \varphi(x)$. Since $\chi_\ell(x)^3 = \det\rho_\ell(x) = \mu^6(\lambda_1\lambda_2)^9$ we conclude that for all $x \in I_\mathfrak{l}^t$ we have

$$\chi_\ell(x)^6 = \mu^{12}(\lambda_1\lambda_2)^{18} = \varphi(x)^{12}(\lambda_2/\lambda_1)^{18} = \varphi(x)^{12},$$

whence for all $x \in I_\mathfrak{l}^t$ there is a fundamental character φ of level 2 such that $\varphi^{6(\ell+1)-12}(x) = 1$. As $|\varphi(I_\mathfrak{l}^t)| = \ell^2 - 1$ for both fundamental characters of level 2 this is absurd for $\ell > 7$. □

Finally, a simple combinatorial argument shows:

Lemma 4.9.14. *For $\ell > 101$ the groups $\mathrm{Sp}_6(\mathbb{F}_\ell)$ and $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$ contain elements of tensor product type with weakly independent eigenvalues (modulo ℓ).*

There are certainly many ways to prove this easy fact, but for the sake of completeness we include a detailed proof:

Proof. Fix a square root $i \in \mathbb{F}_{\ell^2}$ of -1 and an element $a \in \mathbb{F}_\ell^\times$ of multiplicative order at least 5. Let Γ be the multiplicative group $\{c+di \mid (c,d) \in \mathbb{F}_\ell^2,\ c^2+d^2=1\}$, which is isomorphic to either $\mathbb{F}_\ell^\times$ or $\ker\left(\mathrm{Norm} : \mathbb{F}_{\ell^2}^\times \to \mathbb{F}_\ell^\times\right)$ according to whether or not -1 is a square modulo ℓ. Notice that if γ is an element of Γ, then the pair (c,d) is uniquely determined by the equations $c+di=\gamma,\ c-di=1/\gamma$. We can then consider the injective group morphism

$$\begin{array}{ccccc} \sigma : & \Gamma & \to & \mathrm{SL}_2(\mathbb{F}_\ell) \otimes \mathrm{SO}_3(\mathbb{F}_\ell) \\ & \gamma = c+di & \mapsto & \sigma_\gamma := \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \otimes \begin{pmatrix} c & d & 0 \\ -d & c & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{array}$$

which, since $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$ has index 2 in $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \mathrm{SO}_3(\mathbb{F}_\ell)$, maps 2Γ into $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$. Since $|\sigma(2\Gamma)| = |2\Gamma| \geq \frac{\ell-1}{2}$, the lemma will follow if we show that the image of σ contains no more than $50 < \frac{\ell-1}{2}$ operators whose eigenvalues are not weakly independent.

It is clear by construction that the eigenvalues of σ_γ are given by the pairwise products of $\Lambda = \{a^{\pm 1}\}$ and $\Psi = \{1, \gamma^{\pm 1}\}$, so σ_γ has weakly independent eigenvalues if and only if all the solutions to the equation $\left(a^{\varepsilon_1}\gamma^{\delta_1}\right)^2 = a^{\varepsilon_2}\gamma^{\delta_2} \cdot a^{\varepsilon_3}\gamma^{\delta_3}$ with $\varepsilon_j \in \{\pm 1\}, \delta_j \in \{0, \pm 1\}$ are given by $\varepsilon_1 = \varepsilon_2 = \varepsilon_3$ and either $\delta_1 = \delta_2 = \delta_3$ or $\delta_1 = 0$ and $\delta_2 = -\delta_3$. Equivalently, σ_γ has weakly independent eigenvalues if and only if the equation $a^m = \gamma^n$ with $m \in \{0, \pm 2, \pm 4\}$ and $n \in \{0,1,2,3,4\}$ has only the trivial solution $m = n = 0$. Notice that (independently of γ) there are no nontrivial solutions with $n = 0$, because $|m|$ is at most 4 while a has order at least 5. On the other hand, for fixed a, for each pair $(m,n) \in \{0, \pm 2, \pm 4\} \times \{1, \ldots, 4\}$ the equation $a^m = \beta^n$ has at most n solutions β, so in total there are at most $5 \times (1+2+3+4) = 50$ triplets (β, m, n) of solutions to the equation $a^m = \beta^n$. In particular, if γ is different from any of these (at most 50) β's, then σ_γ has weakly independent eigenvalues, and by what we already remarked this finishes the proof. □

4.9.2 Chebotarev bounds

For the proof of theorem 4.1.4 we need one last ingredient, namely an effective version of the Chebotarev density theorem. Lagarias and Odlyzko proved such a result in [53], but their estimate involved a non-explicit constant (which was however effectively computable in principle); their bound was subsequently improved by Œsterlé, who also computed the constant (cf. [92] and [117, §2.5]). To state Œsterlé's result we fix some notation. We let as usual K be a number field, and denote by Δ_K its absolute discriminant; we also write S for a finite subset of Ω_K (the set of finite places of K). To simplify the formulas that follow it is also useful to introduce the function

$$\Delta^*(K,S,N) := |\Delta_K|^N \left(N \cdot \prod_{v \in S} p_v^{1-1/N}\right)^{N \cdot [K:\mathbb{Q}]},$$

where N is a positive integer, and express the bounds we obtain in terms of the quantity

$$B(K,S,N) = 70 \cdot (\log \Delta^*(K,S,N))^2.$$

Theorem 4.9.15. *(Effective Chebotarev under GRH, [92]) Assume the Generalized Riemann Hypothesis. Let L/K be a Galois extension of number fields of degree at most N and let S be a set of finite places of K containing the ones the ramify in L. For every conjugacy class C of $\mathrm{Gal}(L/K)$ there is a place v of K satisfying:*

1. *v is of degree 1 over $\mathbb{Q}$ and does not belong to S;*
2. *the image of Fr_v in $\mathrm{Gal}(L/K)$ lies in C;*
3. *$p_v \leq B(K,S,N)$.*

Remark 4.9.16. Lagarias and Odlyzko also proved a version of theorem 4.9.15 which does not depend on the Generalized Riemann Hypothesis: more precisely, they showed that the same conclusion holds at the cost of replacing $B(K,S,N)$ by $\Delta^*(K,S,N)^c$, where c is an absolute and effectively computable constant. Unpublished work of Winckler [144] shows that one can take $c = 27175010$.

We can finally prove theorem 4.1.4, whose statement we reproduce here for the reader's convenience:

Theorem 4.9.17. *(Theorem 4.1.4) Let A/K be an abelian variety of dimension 3 such that $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. Denote by $\mathcal{N}^0_{A/K}$ the naive conductor of A/K, that is, the product of the prime ideals of $\mathcal{O}_K$ at which A has bad reduction, and suppose that $A[7]$ is defined over K.*

- *Assume the Generalized Riemann Hypothesis: then the equality $G_{\ell^\infty} = \mathrm{GSp}_6(\mathbb{Z}_\ell)$ holds for every prime ℓ unramified in K and strictly larger than $(2q)^{48}$, where*
$$q = b(A^2/K;3)^8 \left(\log|\Delta_{K/\mathbb{Q}}| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right)^2.$$
- *Unconditionally, the same conclusion holds with*
$$q = \exp\left(cb(A^2/K;3)^8 \left(\log|\Delta_K| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right)^2\right),$$
where c is an absolute, effectively computable constant.

Proof. Let ℓ_0 be the smallest prime larger than $b(A^2/K;3)^{1/6}$; by Bertrand's postulate we have $\ell_0 \leq 2b(A^2/K;3)^{1/6}$. Let L denote the field $K(A[\ell_0])$. By construction the Galois group $\mathrm{Gal}\,(L/K)$ is just G_{ℓ_0}, and by proposition 4.9.12 we know that G_{ℓ_0} contains $\mathrm{SL}_2(\mathbb{F}_\ell) \otimes \Omega_3(\mathbb{F}_\ell)$ and hence, by lemma 4.9.14, an operator of tensor product type with weakly independent eigenvalues. Let C be the conjugacy class of this operator and set

$$S = \left\{v \in \Omega_K \;\middle|\; p_v \leq (2g)^2 = 36 \text{ or } A \text{ has bad reduction at } v\right\} \cup \left\{v \in \Omega_K \;\middle|\; p_v = \ell_0\right\}$$

and $N = [L:K]$. Clearly $N \leq |\mathrm{GSp}_6(\mathbb{F}_{\ell_0})| < \ell_0^{22} \leq 2^{22} b(A^2/K;3)^{11/3}$ and

$$\begin{aligned}\log\left(\prod_{v\in S} p_v\right) &\leq \log\left(\ell_0^{[K:\mathbb{Q}]} \cdot \prod_{p<37} p^{[K:\mathbb{Q}]} \cdot \prod_{v \text{ of bad reduction}} p_v\right)\\ &\leq \log N_{K/\mathbb{Q}}(\mathcal{N}^0_{A/K}) + [K:\mathbb{Q}]\,(26.1 + \log\ell_0)\\ &< \log N_{K/\mathbb{Q}}(\mathcal{N}^0_{A/K}) + \frac{1}{3}[K:\mathbb{Q}]\log b(A^2/K;3).\end{aligned}$$

We obtain a (rough) bound on $\Delta^*(K,S,N)$ of the form

$$\begin{aligned}\log\Delta^*(K,S,N) &\leq N\,(\log|\Delta_K| + [K:\mathbb{Q}]\log N +\\ &\qquad + [K:\mathbb{Q}](\log N_{K/\mathbb{Q}}(\mathcal{N}^0_{A/K}) + \frac{1}{3}[K:\mathbb{Q}]\log b(A^2/K;3))\Big)\\ &\leq \frac{1}{\sqrt{70}} b(A^2/K;3)^4 \left(\log|\Delta_K| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)\right),\end{aligned}$$

where on the last line we have used the fact (a deep theorem of Fontaine and Abrashkin) that there are no abelian varieties over $\mathbb{Q}$ having good reduction everywhere, and therefore the term $\log|\Delta_K| + \log N_{K/\mathbb{Q}}\left(\mathcal{N}^0_{A/K}\right)$ is always at least $\log 2$. We now see from theorem 4.9.15 that there exists a place v of K of degree one, satisfying

$$\max\{(2g)^2, \ell_0\} < p_v < 70\,(\log \Delta^*(K,S,N))^2 = q$$

and such that Fr_v maps to the conjugacy class C in $\mathrm{Gal}(L/K) = G_{\ell_0}$. By corollary 4.9.11, Fr_v is not of tensor product type, and by construction A has good reduction at v (recall that $v \notin S$). In particular, we can use this place v to apply theorem 4.1.3 (cf. remark 4.8.1), and the conclusion follows because $(2q)^{48}$ is much larger than either $b(A^2/K;3)^{1/6}$ or $b(A/K;3)$.
Finally, if we do not assume the Generalized Riemann Hypothesis, we get the desired conclusion by applying the unconditional version of the effective Chebotarev theorem, cf. remark 4.9.16. □

Remark 4.9.18. The assumption that $A[7]$ is defined over K is not a serious restriction. Let A/K_0 be any abelian threefold with absolutely trivial endomorphism ring and let K be the field $K_0(A[7])$. Clearly if for some prime ℓ the representation $\rho_\ell^{(K)} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GSp}(A[\ell])$ is surjective, then the same is true for the representation $\rho_\ell^{(K_0)} : \mathrm{Gal}(\overline{K_0}/K_0) \to \mathrm{GSp}(A[\ell])$, so it suffices to give an effective bound ℓ_0 such that $\rho_\ell^{(K)}$ is surjective for $\ell > \ell_0$. Let $S_0 \subseteq \Omega_{K_0}$ be the set of places of bad reduction of A. The degree N of the extension K/K_0 is bounded by $N := |\,\mathrm{GL}_6(\mathbb{F}_7)|$, and it ramifies at most at the places of S_0 and at those of characteristic 7; set $S = S_0 \cup \left\{v \in \Omega_{K_0} \;\middle|\; p_v = 7\right\}$. It follows from [117, Proposition 5] that

$$|\Delta_K| \le \Delta^*(K_0,S,N) < \Delta_{K_0}^N \cdot \left(7^{[K:\mathbb{Q}]}N\right)^{N[K:\mathbb{Q}]} \cdot \left(N_{K_0/\mathbb{Q}}\mathcal{N}^0_{A/K_0}\right)^{N[K:\mathbb{Q}]}.$$

We can then apply theorem 4.9.17 to A/K to get an effective bound ℓ_0 as above, without needing $A[7]$ to be defined over K.

4.10 Class-$\mathcal{S}$ subgroups of Lie type

In view of the result of proposition 4.6.4 we are interested in the question of whether, for a fixed value of n, the group $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ actually contains any class-$\mathcal{S}$ subgroup with simple socle of Lie type (in characteristic ℓ). We have already remarked in section 4.5 that $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ contains maximal class-$\mathcal{S}$ subgroups with socle $\mathrm{PSL}_2(\mathbb{F}_\ell)$ for all n and almost all ℓ; our purpose is to show that in fact, for most n's and ℓ's, *all* the maximal class-$\mathcal{S}$ subgroups of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ of Lie type have socle $\mathrm{PSL}_2(\mathbb{F}_\ell)$:

Theorem 4.10.1. *Set*

$$\mathcal{E} = \left\{ n \in \mathbb{N}, n \text{ odd} \;\middle|\; \begin{array}{c} \text{there exist a prime } \ell > \frac{1}{2}(2n+1)^{12n} \text{ and a maximal} \\ \text{class-}\mathcal{S} \text{ subgroup } G \text{ of } \mathrm{GSp}_{2n}(\mathbb{F}_\ell) \text{ such that } \mathrm{soc}(\mathbb{P}G) \\ \text{is of Lie type in characteristic } \ell \text{ and } \mathrm{soc}(\mathbb{P}\mathrm{G}) \neq \mathrm{PSL}_2(\mathbb{F}_\ell) \end{array} \right\}$$

and let $e(x) = \left|\left\{n \in \mathcal{E} \;\middle|\; n \le x\right\}\right|$ *be the associated counting function. Then for all* $\varepsilon > 0$ *we have* $e(x) = O(x^{2/3+\varepsilon})$*; in particular,* $\mathcal{E}$ *has density zero.*

The proof of this result will take a rather lengthy detour through representation theory: in the next few sections we shall show how to turn the problem at hand into a question about algebraic groups

in positive characteristic, and subsequently reduce this question to a statement about algebraic groups in characteristic 0, which can then be handled by the methods of [32]. From now on, we assume $\ell \neq 2, 3$, so as to avoid the pathologies associated with the finite Suzuki and Ree groups.

4.10.1 Preliminaries on algebraic groups and root systems

Let G be a simple, simply connected algebraic group of rank r over an algebraically closed field. We fix a maximal torus T of G and write $\Lambda \cong \mathbb{Z}^r$ for its character group and $\{\alpha_1, \ldots, \alpha_r\}$ for its simple roots. The vector space $\Lambda \otimes \mathbb{R}$ is in a natural way an Euclidean space, and we write $(\cdot, \cdot)$ for its inner product.

If α is an element of Λ (in particular, if it is a root) we write $\alpha^\vee$ for $\frac{2\alpha}{(\alpha,\alpha)}$, and define the **fundamental weights** $\omega_1, \ldots, \omega_r$ as being the dual basis of $\alpha_i^\vee$ with respect to $(\cdot, \cdot)$. By definition, they satisfy $\left(\omega_i, \alpha_j^\vee\right) = \delta_{ij}$, and they are a $\mathbb{Z}$-basis of Λ (this comes from the fact that G is simply connected). It is also convenient to introduce the map $\langle \cdot, \cdot \rangle : \Lambda \times \Lambda \to \mathbb{Z}$ given by

$$\langle \lambda, \alpha \rangle := \left(\lambda, \alpha^\vee\right) = \frac{2(\lambda, \alpha)}{(\alpha, \alpha)},$$

which allows us to recast the duality between fundamental weights and simple roots in the compact form $\langle \omega_i, \alpha_j \rangle = \delta_{ij}$. Notice that we take the convention that $\langle \cdot, \cdot \rangle$ be linear in its first argument. A weight $\lambda \in \Lambda$ will be said to be **dominant** if $\langle \lambda, \alpha_i \rangle \geq 0$ for all $i = 1, \ldots, r$; equivalently, if it is an integral combination of the fundamental weights ω_i with non-negative coefficients. We denote Λ^+ the cone of dominant weights. We can introduce a partial ordering (both on Λ and on Λ^+) by declaring that a weight λ is larger than a weight μ (in symbols, $\lambda \succ \mu$) if and only if $\lambda - \mu$ can be written as a sum of simple roots with *non-negative* coefficients.

We also write Δ for the set of all roots of G, and Δ^+ for the subset of positive roots, i.e. those that can be written as integral linear combinations of the α_i's with non-negative coefficients; we have $|\Delta| = 2|\Delta^+|$. We define the **Weyl vector** δ by the formula $\delta = \frac{1}{2}\sum_{\alpha \in \Delta^+} \alpha$, and recall ([42, §13.3, Lemma A]) that $\delta = \sum_{i=1}^r \omega_i$.

The **Coxeter number** of G is defined to be the ratio $h := \frac{|\Delta|}{r} = \frac{2|\Delta^+|}{r}$. By the classification of simple root systems it is known that h does not exceed $4r$ (and is in fact at most $2r$ as long as $r \geq 9$).

The **Cartan matrix** of a root system (relative to a given choice of simple roots) is the $r \times r$ matrix whose (i, j)-th entry is given by $C_{ij} = \langle \alpha_i, \alpha_j \rangle$. Writing a simple root α_i as a combination of the fundamental weights, $\alpha_i = \sum_{j=1}^r b_j \omega_j$, and applying the linear map $\langle \cdot, \alpha_k \rangle$ to both sides of this equation we obtain $C_{ik} = b_k$, so the Cartan matrix is the base-change matrix expressing the simple roots in terms of the fundamental weights. Moreover, C enjoys the following property, which can be gleaned from a direct inspection of tables I through IX of Bourbaki [17]:

Lemma 4.10.2. *The matrix $C - 2\,\mathrm{Id}$ has non-positive entries and its diagonal coefficients vanish.*

Finally, recall that the **Weyl group** of G, denoted by $W(G)$, is the subgroup of $\mathrm{GL}(\Lambda \otimes \mathbb{R})$ generated by the reflections along the simple roots α_i, and that the same definition can also be used to introduce a notion of Weyl group for not necessarily irreducible root systems and for not necessarily connected Dynkin diagrams. If Δ (resp. D) is a root system (resp. the associated Dynkin diagram) we write $W(\Delta) = W(D)$ for the corresponding Weyl group.

We conclude this section of preliminaries with a simple lemma which is certainly well-known to experts, but for which we could not find any reference in the literature:

Lemma 4.10.3. *Suppose G is of rank r and let $\lambda \in \Lambda$ be a nonzero weight. The orbit of λ under the Weyl group of G contains at least $r+1$ distinct weights.*

Proof. Let D be the Dynkin diagram associated with the root system of G. By the orbit-stabilizer lemma it is enough to show that the stabilizer of λ has index at least $r+1$ in $W(D)$. Since every weight is $W(D)$-conjugated to a dominant weight, there is no loss of generality in assuming that λ is dominant. In this case, the stabilizer of λ is known to be generated by those reflections s_α along simple roots such that $s_\alpha \lambda = \lambda$ ([42, §10.3B]). Since the stabilizer of λ is clearly not the full Weyl group $W(D)$, there is at least one simple root β whose associated reflection does not stabilize λ. The stabilizer of λ is then identified to a subgroup of the group generated by s_α for all simple roots $\alpha \neq \beta$; notice that the group generated by $\left\{ s_\alpha \;\middle|\; \alpha \text{ a simple root}, \alpha \neq \beta \right\}$ is isomorphic to the Weyl group of the Dynkin diagram obtained from D by erasing the node corresponding to β. We thus obtain the following procedure for determining a lower bound on the index of $\mathrm{Stab}(\lambda)$ in $W(D)$: we consider the Dynkin diagram D and all the (quite possibly non-connected) diagrams $D_1, \ldots, D_r$ which we can obtain from D by erasing exactly one node. We then compute the Weyl groups $W(D_i)$ associated with each of these diagrams and the indices $|W(D)/W(D_i)|$: the smallest such index is a lower bound for the index $|W(D)/\mathrm{Stab}(\lambda)|$. The lemma now follows from a straightforward, if somewhat tedious, examination of the connected Dynkin diagrams and of table 4.1. As an example, let us do this for root systems of type A_r, which give the smallest possible index. Removing the i-th node ($i = 1, \ldots, r$) from the Dynkin diagram for A_r leads to the Dynkin diagram for the root system $A_{i-1} \times A_{r-i}$, where by $A_0 \times A_{r-1}$ and $A_{r-1} \times A_0$ we simply mean A_{r-1}. The Weyl group of this root system is $S_i \times S_{r-i+1}$, whose index in the Weyl group of A_r is $\frac{(r+1)!}{(i)!(r-i+1)!} = \binom{r+1}{i} \geq r+1$. □

Root system	Order of the Weyl group
A_n	$(n+1)!$
B_n	$2^n n!$
C_n	$2^n n!$
D_n	$2^{n-1} n!$
E_6	$72 \cdot 6!$
E_7	$72 \cdot 8!$
E_8	$192 \cdot 10!$
F_4	1152
G_2	12

TABLE 4.1: Order of Weyl groups

4.10.2 Representation theory of finite simple groups of Lie type

This paragraph is essentially taken from [65], which will be our main reference for this section; further information can be found in [20], Chapter 1 (especially sections 1.17-1.19). Let $\tilde{G}$ be a finite twisted or non-twisted simple Chevalley group in characteristic $\ell \neq 2, 3$ (that is, a finite simple

group of Lie type of characteristic different from 2 and 3; in particular, not a Suzuki or a Ree group). We shall describe shortly the main algebraic data associated with $\tilde{G}$, but before doing so we need to define Frobenius maps:

Definition 4.10.4. Let k be an algebraically closed field of characteristic $\ell > 0$, and let $q = \ell^e$ (where e is a positive integer). The q-Frobenius map of $\mathrm{GL}_n(k)$, denoted F_q, is the automorphism of $\mathrm{GL}_n(k)$ that raises all coefficients of a matrix to the q-th power. Let G be a linear algebraic group over k. A **standard Frobenius map** is a group morphism $F : G(k) \to G(k)$ such that, for some embedding $\iota : G(k) \hookrightarrow \mathrm{GL}_n(k)$ and for some $q = \ell^e$, the identity $\iota(F(g)) = F_q(\iota(g))$ holds for every $g \in G(k)$. Finally, a group morphism $G(k) \to G(k)$ is a **Frobenius map** (or endomorphism) if some power of it is a standard Frobenius map.

It is known that to a group $\tilde{G}$ as above we can attach a connected reductive simple algebraic group G over $\overline{\mathbb{F}_\ell}$ of simply connected type and a Frobenius endomorphism F of G with the following property: denoting by G^F the group $\left\{g \in G(\overline{\mathbb{F}_\ell}) \;\middle|\; F(g) = g\right\}$ of fixed points of F, and by Z the center of G^F, we have $\tilde{G} \cong G^F/Z$. Furthermore, G^F is the universal covering group (also known as the universal perfect central extension) of $\tilde{G}$, see [29] and the references therein.

Remark 4.10.5. It is further known that the Frobenius endomorphism F is completely characterised by the choice of an automorphism of the Dynkin diagram of G together with a real number q which, in our setting, is an integral power of ℓ. We include this number q among the data associated with $\tilde{G}$; it will appear for example in the statements of theorem 4.10.6 and in the proof of lemma 4.10.24.

In this situation, we shall call G the algebraic group associated with $\tilde{G}$, and we shall indifferently speak of the rank of $\tilde{G}$, of G^F, or of G; likewise, we shall say that $\tilde{G}, G^F$, or G, is of type A_r (resp. $B_r, C_r, \ldots$) if the root system of G is.
Our interest in this construction comes from the fact that projective representations of $\tilde{G}$ in characteristic ℓ are the same as linear representations of G^F in characteristic ℓ ([131, pp. 76-77, items (ix) and (x)]), which in turn can be constructed by restricting algebraic representations of the algebraic group G to G^F, as we now describe. Let G be of rank r, denote by Λ^+ the cone of its dominant weights (with respect to a given maximal torus), and write $\omega_1, \ldots, \omega_r$ for the fundamental ones; for any given dominant weight $\lambda \in \Lambda^+$, the irreducible $\overline{\mathbb{F}_\ell}[G]$-module with highest weight λ will be denoted by $L(\lambda)$. The relationship between representations of G^F and algebraic representations of G is nicely described by the following theorem of Steinberg:

Theorem 4.10.6. *(Steinberg [130]) Let G, G^F and q be as above (with the restriction that the characteristic be different from $2, 3$). Define*

$$\Lambda_q = \left\{a_1\omega_1 + \cdots + a_r\omega_r \;\middle|\; 0 \leq a_i \leq q-1 \text{ for } 1 \leq i \leq r\right\}.$$

The restrictions of the G-modules $L(\lambda)$ with $\lambda \in \Lambda_q$ to G^F form a set of pairwise inequivalent representatives of all equivalence classes of irreducible $\overline{\mathbb{F}_\ell}[G^F]$-modules.

4.10.3 Some structure theorems

In this section we recall further results that describe more finely the structure of the simple modules $L(\lambda)$. It is useful to introduce the notion of (m-)restricted weights:

Definition 4.10.7. Let G, G^F be as above and m be a positive integer. A dominant weight $\lambda = a_1\omega_1 + \ldots + a_r\omega_r \in \Lambda^+$ is said to be m-restricted if for every $i = 1, \ldots, r$ we have $0 \leq a_i \leq m-1$.

Definition 4.10.8. Let F be an automorphism of a group $\tilde{G}$ and $\rho : \tilde{G} \to \mathrm{Aut}(V)$ be a representation of $\tilde{G}$. The **twist** of ρ by F is the representation ${}^F\rho$ given by ${}^F\rho(g) = \rho(F(g))$ for all $g \in G$. Note that twisting the representation does not change its image, nor its dimension.

The field automorphism $x \mapsto x^\ell$ of $\overline{\mathbb{F}_\ell}$ can be used to construct a canonical endomorphism of the algebraic group G, called the 'standard Frobenius map' and denoted by F_0 ([43, §2.7]).
The following theorem elucidates the importance of ℓ-restricted weights and their interactions with Frobenius twists:

Theorem 4.10.9. *(Steinberg's twisted tensor product theorem [130]) If L is a G-module, let $L^{(i)}$ be the module obtained by twisting the G-action on L by F_0^i. If $\lambda_0, \ldots, \lambda_m$ are ℓ-restricted weights, then*

$$L(\lambda_0 + \ell\lambda_1 + \ldots + \ell^m\lambda_m) \cong L(\lambda_0) \otimes L(\lambda_1)^{(1)} \otimes \cdots \otimes L(\lambda_m)^{(m)}.$$

Theorems 4.10.6 and 4.10.9 are all we need to describe representations over $\overline{\mathbb{F}_\ell}$. However, to deal with groups with socle $\mathrm{PSL}_2(q)$, where q is a power of ℓ different from ℓ, it is not enough to work over $\overline{\mathbb{F}_\ell}$, but we shall need to know when a representation over $\overline{\mathbb{F}_\ell}$ can be defined over a smaller field. We make this notion more precise in the following definition:

Definition 4.10.10. Let $\tilde{G}$ be a finite group, K a field, and $\rho : \tilde{G} \to \mathrm{GL}_n(K)$ a representation of $\tilde{G}$ over K. We say that ρ can be defined over a field $k \subseteq K$ if there exists a representation $\rho_k : \tilde{G} \to \mathrm{GL}_n(k)$ such that the representation

$$\tilde{G} \xrightarrow{\rho_k} \mathrm{GL}_n(k) \hookrightarrow \mathrm{GL}_n(K)$$

is isomorphic to ρ over K.

The fields of definition of modular representations of finite groups of Lie type are very well understood (cf. [19, Theorem 5.1.13]). Here we just need the simplest case, namely a criterion to decide whether a representation can be defined over $\mathbb{F}_\ell$:

Proposition 4.10.11. *Let $\ell \neq 2, 3$. Write the number q associated with G^F (cf. remark 4.10.5) as ℓ^e. Let M be an irreducible module for G^F, and write M as a tensor product $\bigotimes_{i=0}^{e-1} M_i^{(i)}$ as in theorem 4.10.9: M can be defined over the field $\mathbb{F}_\ell$ if and only if $M_i \cong M_j$ for all i, j.*

Proof. This follows at once from the proof of [19, Theorem 5.1.13]. More specifically, by [19, Corollary 1.8.14] M can be defined over $\mathbb{F}_\ell$ if and only if it is stabilized by the Frobenius automorphism F_0, and on the other hand by definition F_0^e is the identity of G^F, so M is isomorphic to $M^{(1)}$ if and only if $\bigotimes_{i=0}^{e-1} M_i^{(i)} \cong \bigotimes_{i=0}^{e-1} M_{i-1}^{(i)}$, where $M_{-1} = M_{e-1}$. Since the representation of theorem 4.10.9 is unique, this implies $M_{i-1} \cong M_i$ for $i = 0, \ldots, e-1$. □

Corollary 4.10.12. *Let $q = \ell^e$ be the invariant attached to G^F, and let M be an absolutely irreducible $\mathbb{F}_\ell$-module for G^F whose dimension n is not a perfect power. Then $e = 1$, that is, $q = \ell$. In particular, this holds if $n \equiv 2 \pmod 4$.*

Proof. Let $\overline{M} := M \otimes \overline{\mathbb{F}_\ell}$ and λ be the associated dominant q-restricted weight. We can write $\lambda = \sum_{i=0}^{e-1} \ell^i \lambda_i$, where each λ_i is ℓ-restricted. By theorem 4.10.9 we have $\overline{M} \cong \bigotimes_{i=0}^{e-1} L(\lambda_i)^{(i)}$, and since by assumption $\overline{M}$ can be descended to $\mathbb{F}_\ell$ the previous proposition gives $L(\lambda_i) \cong L(\lambda_j)$ for all i, j. It follows that $n = \dim(L_\lambda) = (\dim L(\lambda_0))^e$, which is incompatible with $e > 1$. Finally notice that no integer $n \equiv 2 \pmod 4$ can be a perfect power, because any power of an even number is divisible by 4. □

4.10.4 Weyl modules

We briefly recall the most basic properties of the so-called Weyl modules; for more information, cf. [43, §3.1]. For any $\lambda \in \Lambda^+$ there is a certain $\mathbb{Z}G$-module $V(\lambda)_\mathbb{Z}$ such that

- the module $L(\lambda)$ is a quotient of $V(\lambda)_\mathbb{Z} \otimes_\mathbb{Z} \overline{\mathbb{F}_\ell}$;
- for a complex, simply connected, simple Lie group $G_\mathbb{C}$ with the same root system as G, the $\mathbb{C}G$-module $V(\lambda)_\mathbb{C} := V(\lambda)_\mathbb{Z} \otimes_\mathbb{Z} \mathbb{C}$ is the unique irreducible module of highest weight λ.

Definition 4.10.13. We call $V(\lambda)_\mathbb{Z} \otimes_\mathbb{Z} \overline{\mathbb{F}_\ell}$ the **Weyl module** associated with λ. It is a $\overline{\mathbb{F}_\ell}[G]$-module which we will denote by $V(\lambda)$.

The celebrated Weyl dimension formula gives the dimension of $V(\lambda)$:

Theorem 4.10.14. *(Weyl) For all dominant weights λ we have*

$$\dim_{\overline{\mathbb{F}_\ell}} V(\lambda) = \dim_\mathbb{C} V(\lambda)_\mathbb{C} = \frac{\prod_{\alpha \in \Delta^+} (\lambda + \delta, \alpha)}{\prod_{\alpha \in \Delta^+} (\lambda, \alpha)},$$

where $\delta = \frac{1}{2} \sum_{\alpha \in \Delta^+} \alpha = \sum_{i=1}^r \omega_i$.

4.10.4.1 Sufficient condition for the equality $V(\lambda) = L(\lambda)$

In general, it can very well happen that $\dim_{\overline{\mathbb{F}_\ell}} L(\lambda)$ is strictly smaller than $\dim_{\overline{\mathbb{F}_\ell}} V(\lambda)$. The following theorem gives interesting information about the action of G^F on Weyl modules, which we shall use to deduce sufficient conditions for $V(\lambda)$ and $L(\lambda)$ to be isomorphic.

Theorem 4.10.15. *(Wong, [147, (2D)], [43, §5.9]) If λ is a q-restricted, dominant weight, the Weyl module $V(\lambda)$ is indecomposable (but not necessarily irreducible) upon restriction to G^F. In particular, it is also indecomposable under the action of G.*

Since $V(\lambda)$ has highest weight λ by construction, $V(\lambda)$ admits a unique G-simple quotient that is the unique irreducible representation of G with highest weight λ; that is to say, $L(\lambda)$ is the unique simple quotient of $V(\lambda)$. We shall now see that, under suitable assumptions on the dimension of $V(\lambda)$ and on ℓ, we must in fact have $V(\lambda) = L(\lambda)$. The key result we need is the following theorem of McNinch (which builds on previous work of Jantzen, [46]).

Theorem 4.10.16. *([76]) Let k be an algebraically closed field of characteristic $\ell \geq 7$, and suppose that the root system of G is not of type A_1. Let furthermore V be a module over $k[G^F]$ such that $\dim_k V \leq 2\ell$: then V is completely reducible.*

Corollary 4.10.17. *Suppose G is not of Lie type A_1. If λ is a dominant and q-restricted weight, ℓ is at least 7, and $\dim V(\lambda) \leq 2\ell$, then $L(\lambda) \cong V(\lambda)$.*

Proof. Notice that an indecomposable and completely reducible module is simple. Hence in particular $V(\lambda)$ is G^F-simple by the combination of the previous theorems, and since $L(\lambda)$ is a simple (nonzero) quotient of $V(\lambda)$ the two must coincide. □

4.10.4.2 The case $V(\lambda) \neq L(\lambda)$

When $L(\lambda)$ does not coincide with $V(\lambda)$ its precise structure is still quite mysterious and forms the subject of a rich body of work. For our applications, however, we shall just need to know that the dimension of $L(\lambda)$ grows reasonably quickly when the coefficients a_i in the representation $\lambda = \sum a_i\omega_i$ go to infinity. To prove such an estimate we shall need the following theorem of Premet:

Theorem 4.10.18. *(Premet, [102]) Let G be a simple, simply connected algebraic group in characteristic ℓ. If the root system of G has different root lengths we assume that $\ell \neq 2$, and if G is of type G_2 we also assume that $\ell \neq 3$. Let λ be an ℓ-restricted dominant weight. The set of weights of the irreducible G-module $L(\lambda)$ is the union of the $W(G)$-orbits of dominant weights μ that satisfy $\mu \prec \lambda$.*

The next lemma provides a lower bound on $\dim L(\lambda)$. The result is almost identical to [33, Lemma 2.3], which is however only stated and proved for root systems of type A_r. As it turns out, a very small modification of the proof given in [33] yields a uniform bound for all root systems.

Lemma 4.10.19. *Let $\lambda = \sum_{i=1}^{r} a_i\omega_i \in \Lambda^+$ be an ℓ-restricted weight. Then*

$$\dim L(\lambda) \geq N(\lambda) := 1 + (r+1)\left\{\prod_{i=1}^{r}\left(\left\lfloor\frac{a_i}{2}\right\rfloor + 1\right) - 1\right\}$$

Proof. Fix r integers $x_1, \ldots, x_r$ with $0 \leq x_i \leq \lfloor\frac{a_i}{2}\rfloor$. Set $\gamma := \sum x_i\alpha_i$ and let C_{ij} be the Cartan matrix of the relevant root system. By lemma 4.10.2, we have $\alpha_i = 2\omega_i - \sum_{j\neq i}|C_{ij}|\omega_j$ since all off-diagonals coefficients of the Cartan matrix are non-positive. It follows that the coefficient of

$$\gamma = \sum_{i=1}^{r} 2x_i\omega_i - \sum_{i=1}^{r}\sum_{j\neq i}|C_{ij}|x_i\omega_j$$

along ω_i, call it b_i, is at most $2x_i \leq a_i$. Hence $\mu := \lambda - \gamma = \sum_{i=1}^{r}(a_i - b_i)\omega_i$ is a linear combination of fundamental weights with non-negative coefficients, hence it is a dominant weight. On the other hand, it is clear that $\lambda \succ \mu$, since $\lambda - \mu = \gamma$ is by construction a combination of simple roots with non-negative coefficients.

There are $\prod_{i=1}^{r}\left(\left\lfloor\frac{a_i}{2}\right\rfloor + 1\right)$ possible choices for the integers x_i, so the module $V(\lambda)$ contains at least $\prod_{i=1}^{r}\left(\left\lfloor\frac{a_i}{2}\right\rfloor + 1\right)$ different dominant weights, at most one of which is the zero weight. Consider now the orbits of the nonzero dominant weights under the Weyl group. Each orbit consists entirely of weights of $V(\lambda)$, and contains exactly one dominant weight. In particular, two orbits do not intersect

(for otherwise we would find two Weyl-conjugated dominant weights); moreover, by lemma 4.10.3 every nontrivial weight has orbit of length at least $r+1$. We have thus found at least

$$1+(r+1)\left\{\prod_{i=1}^{r}\left(\left\lfloor\frac{a_i}{2}\right\rfloor+1\right)-1\right\}=N(\lambda)$$

distinct weights in $V(\lambda)$. Premet's theorem 4.10.18 implies that these weights all appear in $L(\lambda)$, which is therefore of dimension at least $N(\lambda)$, as claimed. □

We derive in particular the following lower bound on $\dim L(\lambda)$:

Proposition 4.10.20. *Let $n\geq 2$ be a positive integer, and suppose that r, the rank of G, satisfies $2\leq r\leq \min\left\{n,\sqrt{6n}\right\}$. If $\lambda=\sum_{i=1}^{r}a_i\omega_i\in\Lambda^+$ is an ℓ-restricted weight such that $\sum_{i=1}^{r}a_i>2n$, then $\dim L(\lambda)>2n$.*

Proof. The previous lemma gives $\dim L(\lambda)\geq N(\lambda)\geq 1+(r+1)\left(\frac{1}{2}\sum_{i=1}^{r}a_i-\frac{r}{2}\right)$, where the second inequality is an equality if all but one of the a_i's are equal to 1, and the remaining one is odd. It is straightforward to check that, for $r\leq n$, the number $1+(r+1)\left(\frac{1}{2}\sum_{i=1}^{r}a_i-\frac{r}{2}\right)$ is not smaller than $2n+1$. □

4.10.5 Lifting to characteristic zero

The purpose of this section is to show that, when the characteristic ℓ is large enough (compared to n), the representation theory of subgroups of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ is equivalent to the representation theory of certain corresponding (algebraic) groups in characteristic zero. In order to do so, we need to ensure that the equality $L(\lambda)=V(\lambda)$ holds for all the λ's of interest, and in view of corollary 4.10.17 it is enough to know that the dimension of $V(\lambda)$ is less than 2ℓ. The following lemma provides an upper bound on the dimension of Weyl modules:

Lemma 4.10.21. *Fix a positive integer n. Consider all simply connected, simple algebraic groups G over $\overline{\mathbb{F}_\ell}$ of rank at least 2 and at most $\min\left\{\sqrt{6n},n\right\}$. For each such G (of rank r), consider the collection of all dominant, ℓ-restricted weights $\lambda=\sum_{i=1}^{r}a_i\omega_i$ such that $\sum_{i=1}^{r}a_i\leq 2n$ and the corresponding Weyl modules $V(\lambda)$. For every such $V(\lambda)$ we have*

$$\dim V(\lambda)\leq(2n+1)^{12n}.$$

Proof. Take a group G (of rank r) and a weight λ as in the statement of the lemma. Notice that any positive root α can be represented as $\alpha=\sum_{j=1}^{r}b_j\alpha_j$, where the b_j are non-negative integers; a simple computation (using the fact that $\delta=\sum_{i=1}^{r}\omega_i$) gives

$$\langle\lambda,\alpha\rangle=\sum_{i=1}^{r}a_ib_i,\qquad\langle\delta,\alpha\rangle=\sum_{j=1}^{r}b_j,$$

so the ratio $\frac{\langle\lambda,\alpha\rangle}{\langle\delta,\alpha\rangle}$ is bounded above by $\max a_i\leq\sum_{i=1}^{r}a_i$. By Weyl's dimension formula we have

$$\begin{aligned}\dim V(\lambda)&=\frac{\prod_{\alpha\in\Delta^+}(\delta+\lambda,\alpha)}{\prod_{\alpha\in\Delta^+}(\delta,\alpha)}\\&=\prod_{\alpha\in\Delta^+}\left(1+\frac{\langle\lambda,\alpha\rangle}{\langle\delta,\alpha\rangle}\right);\end{aligned}$$

combining this formula, the arithmetic-geometric inequality and the bound $\frac{\langle\lambda,\alpha\rangle}{\langle\delta,\alpha\rangle} \leq \sum_{i=1}^{r} a_i \leq 2n$ we deduce

$$\dim V(\lambda) \leq \left(\frac{\sum_{\alpha\in\Delta^+}(1+\sum a_i)}{|\Delta^+|}\right)^{|\Delta^+|} \leq (2n+1)^{|\Delta^+|}.$$

Finally, since the Coxeter number h does not exceed $4r$, we have $|\Delta^+| = \frac{rh}{2} \leq 2r^2 \leq 12n$, and thus $\dim V(\lambda) \leq (2n+1)^{12n}$ as claimed. □

The following proposition gives the desired lift to characteristic zero, assuming that ℓ is large enough with respect to n:

Proposition 4.10.22. *Let n be an odd integer and ℓ a prime not smaller than $\frac{1}{2}(2n+1)^{12n}$. Then for all groups of the form G^F (where $\operatorname{rank}(G) \leq \min\{\sqrt{6n}, n\}$), and for all absolutely irreducible representations V of G^F over $\mathbb{F}_\ell$ of dimension $2n$, there exist*

- *a simple, simply connected, complex Lie group $G_{\mathbb{C}}$ with the same Lie algebra (hence in particular the same rank) as G;*
- *a complex, irreducible representation $V_{\mathbb{C}}$ of $G_{\mathbb{C}}$ such that $\dim_{\mathbb{C}} V_{\mathbb{C}} = 2n$.*

Proof. Let G^F, G and V be as in the statement, and let r be the rank of G. By corollary 4.10.12 we have $q = \ell$, and by theorem 4.10.6 $V \otimes \overline{\mathbb{F}_\ell}$ is of the form $L(\lambda)$ for a q-restricted (hence ℓ-restricted) weight λ. Write λ as $\sum_{i=1}^{r} a_i\omega_i$, and notice that $\sum_{i=1}^{r} a_i \leq 2n$, for otherwise $\dim V = \dim L(\lambda) > 2n$ by proposition 4.10.20, a contradiction. Lemma 4.10.21 then gives $\dim V(\lambda) \leq (2n+1)^{12n} \leq 2\ell$, which by corollary 4.10.17 implies $L(\lambda) \cong V(\lambda)$. Now if $G_{\mathbb{C}}$ is the unique simple, simply connected, complex Lie group with the same root system as G, then $V(\lambda)_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{C}$ is exactly the (irreducible, complex) representation associated with the dominant weight $\sum_{i=1}^{r} a_i\omega_i$ of $G_{\mathbb{C}}$, and the result follows. □

4.10.6 Zero-density estimate in characteristic zero

We have now essentially turned our problem into a question about the representation theory of certain complex Lie groups, for which we have the following zero-density estimate:

Proposition 4.10.23. *Let C be any positive real number. Set*

$$\mathcal{E}_C = \left\{ n \in \mathbb{N} \;\middle|\; \begin{array}{c} \text{there exists a simple, simply connected, complex Lie group } G \\ \text{of rank } r, \text{ with } 2 \leq r \leq C\sqrt{n}, \\ \text{admitting an irreducible representation on } \mathbb{C}^n \end{array} \right\}$$

and let $e_C(x) = \left|\{n \in \mathcal{E}_C \mid n \leq x\}\right|$ be the associated counting function. Then for all $\varepsilon > 0$ we have $e_C(x) = O(x^{2/3+\varepsilon})$; in particular, $\mathcal{E}$ has density zero.

Proof. Fix $\varepsilon > 0$. Let $\mathcal{L}$ be the collection of all (isomorphism classes of) complex, simple, simply connected Lie groups of rank at least 2, and let $\mathcal{L}_x$ be the subset of those having rank at most $C\sqrt{x}$. All the groups in $\mathcal{L}$ have Coxeter number at least 3: the only Lie algebra with Coxeter number 2 is A_1, which we have excluded. Also note that $|\mathcal{L}_x| = O(x^{1/2})$: there are at most 5 Lie algebras of any fixed rank.

Following [56], for a complex, simple, simply connected Lie group G we denote by $R_x(G)$ the number of isomorphism classes of irreducible representations of G of dimension at most x. It follows from [56, Theorem 5.1] that for every $G \in \mathcal{L}$ we have $R_x(G) = O(x^{2/3+\varepsilon})$.
Furthermore, by [32, Corollary 3] we know that there exists a finite subset Σ_ε of $\mathcal{L}$ (depending on ε), such that, for all $G \in \mathcal{L} \setminus \Sigma_\varepsilon$, the inequality $R_x(G) \leq x^\varepsilon$ holds for every $x \geq 1$. Note that in fact [32] deals with compact Lie groups, but as it is well known every simple complex simply connected Lie group admits a unique compact real form which has the same representation theory as the complex group, so the result holds in our setting as well. It is now clear that

$$\begin{aligned} e_C(x) \leq \sum_{G \in \mathcal{L}_x} R_x(G) &= \sum_{G \in \mathcal{L}_x \setminus \Sigma_\varepsilon} R_x(G) + \sum_{G \in \Sigma_\varepsilon} R_x(G) \\ &\leq \sum_{G \in \mathcal{L}_x \setminus \Sigma_\varepsilon} x^\varepsilon + \sum_{G \in \Sigma_\varepsilon} O(x^{2/3+\varepsilon}) \\ &= O(x^{1/2+\varepsilon}) + O(x^{2/3+\varepsilon}) = O(x^{2/3+\varepsilon}). \end{aligned}$$

□

4.10.7 Order estimates

We now invoke simple order estimates to show that if the finite simple group of Lie type H appears as a class-$\mathcal{S}$ subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$, then its rank cannot exceed $\sqrt{6n}$.

Lemma 4.10.24. *Let L be a finite simple group of Lie type in characteristic $\ell \neq 2, 3$ and r be its rank (i.e. the rank of the corresponding algebraic group): we have $|L| \geq \ell^{r^2}$.*

Proof. The group in question is characterized by a number q (a integral power of ℓ) and by the family to which it belongs. For most families of simple Lie groups, the claim is easy to check by direct inspection of the explicit formulas for the orders, so let us only check families $A_r(q)$ and ${}^2A_r(q^2)$, which are arguably the least trivial ones. In the two cases, the order is given by

$$\begin{aligned} \frac{q^{r(r+1)/2}}{(r+1, q-\varepsilon)} \prod_{i=1}^{r} (q^{i+1} - \varepsilon^{i+1}) &\geq \frac{q^{r(r+1)/2}}{q(q+1)} q^{(r+1)(r+2)/2} \prod_{i=1}^{r} (1 - (\varepsilon q)^{-i-1}) \\ &\geq \frac{q^{(r+1)^2}}{q(q+1)} \prod_{i=1}^{\infty} \left(1 - \frac{1}{q^{i+1}}\right), \end{aligned}$$

where $\varepsilon = +1$ for $A_r(q)$ and $\varepsilon = -1$ for ${}^2A_r(q^2)$. On the other hand,

$$\log \prod_{i=1}^{\infty} (1 - q^{-i-1}) = \sum_{i=1}^{\infty} \log\left(1 - q^{-i-1}\right) \geq \sum_{i=1}^{\infty} -2q^{-i-1} = -\frac{2}{q(q-1)} \geq -\frac{1}{10}$$

The order of the group in question is thus at least $\exp(-1/10) \dfrac{q}{q(q+1)} q^{2r} \cdot q^{r^2} > q^{r^2} \geq \ell^{r^2}$ as claimed. □

We now compare this lower bound with the following upper bound due to Liebeck:

Theorem 4.10.25. *([63, Main theorem]) Let n be a positive integer and H be a class-$\mathcal{S}$ subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$. The order of $\mathbb{P}H$ is strictly smaller than $\max\left\{\ell^{6n}, (2n+2)!\right\}$.*

Since $\ell^{6n} > (2n+2)!$ for $\ell > 2n+2$ we also have:

Corollary 4.10.26. *In the situation of the previous theorem, suppose $\ell > 2n+2$. Then the order of $\mathbb{P}H$ is strictly smaller than ℓ^{6n}.*

Corollary 4.10.27. *Let H be a class-$\mathcal{S}$ subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$, with $\ell > 2n+2$. Suppose $\mathrm{soc}(\mathbb{P}H)$ is simple of Lie type in characteristic ℓ: then the rank of $\mathrm{soc}(\mathbb{P}H)$ is less than $\sqrt{6n}$.*

Proof. Indeed, if r denotes the rank of $\mathrm{soc}(\mathbb{P}H)$ we have $|\mathrm{soc}(\mathbb{P}H)| \geq \ell^{r^2}$ by lemma 4.10.24 and $|\mathrm{soc}(\mathbb{P}H)| < \ell^{6n}$ by corollary 4.10.26. □

4.10.8 Conclusion in positive characteristic

We are finally ready for the proof of theorem 4.10.1:

Proof. We can assume without loss of generality that $n \geq 6$, so that $\min\{n, \sqrt{6n}\} = \sqrt{6n}$. We claim that the set $\{2n \mid n \in \mathcal{E}, n \geq 6\}$ is contained in the set $\mathcal{E}_C$ of proposition 4.10.23 for $C = \sqrt{3}$. Indeed let $n \geq 6$ be an element of $\mathcal{E}$: then we can find

- a prime $\ell > \frac{1}{2}(2n+1)^{12n}$;
- a finite group of Lie type in characteristic ℓ, call it G^F, different from $\mathrm{SL}_2(\mathbb{F}_\ell)$;
- and an absolutely irreducible representation of G^F in characteristic ℓ of degree $2n$.

Note that by corollary 4.10.12 we cannot have $G^F = \mathrm{SL}_2(q)$ for $q = \ell^e$, $e > 1$. Let G be the simple, simply connected algebraic group associated with G^F as in section 4.10.2. By corollary 4.10.27 we have $\mathrm{rank}(G) < \sqrt{6n}$, and by what we just remarked we have $\mathrm{rank}(G) \geq 2$.
We are now in the situation of proposition 4.10.22, so we find a simple, simply connected, complex Lie group $G_{\mathbb{C}}$, of rank lying in the interval $[2, \sqrt{3} \cdot \sqrt{2n}]$, admitting an irreducible representation on $\mathbb{C}^{2n}$. By definition, this means that $2n \in \mathcal{E}_C$ (for $C = \sqrt{3}$). In particular, the counting function $e(x)$ satisfies $e(x) \leq e_C(2x)$, and therefore it is $O(x^{2/3+\varepsilon})$ for any positive ε by proposition 4.10.23. □

4.11 Explicit determination of the small exceptional dimensions

In this section we determine all odd $n \leq 100$ with the following property: for at least one prime $\ell > 13$, the group $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ contains a class-$\mathcal{S}$ subgroup H such that $\mathrm{soc}(\mathbb{P}H)$ is of Lie type in characteristic ℓ and different from $\mathrm{PSL}_2(\mathbb{F}_\ell)$. In order to carry out explicit calculations, we need to fix our convention for the simple roots, and since we are going to rely on the tables of [65] we adopt the same numbering as in that paper, which we summarize in table 4.2 (note that this convention does not agree with that of [17]).
We shall need some information about the duality properties of our representations; recall that the Frobenius-Schur indicator of an irreducible representation is $+1$ if that representation is orthogonal, -1 if it is symplectic, and 0 if it is not self-dual. Regarding the Frobenius-Schur indicator of the modular representations we are interested in we have the following result of Steinberg ([131, Lemmas 78 and 79], but cf. also [65, §6.3]):

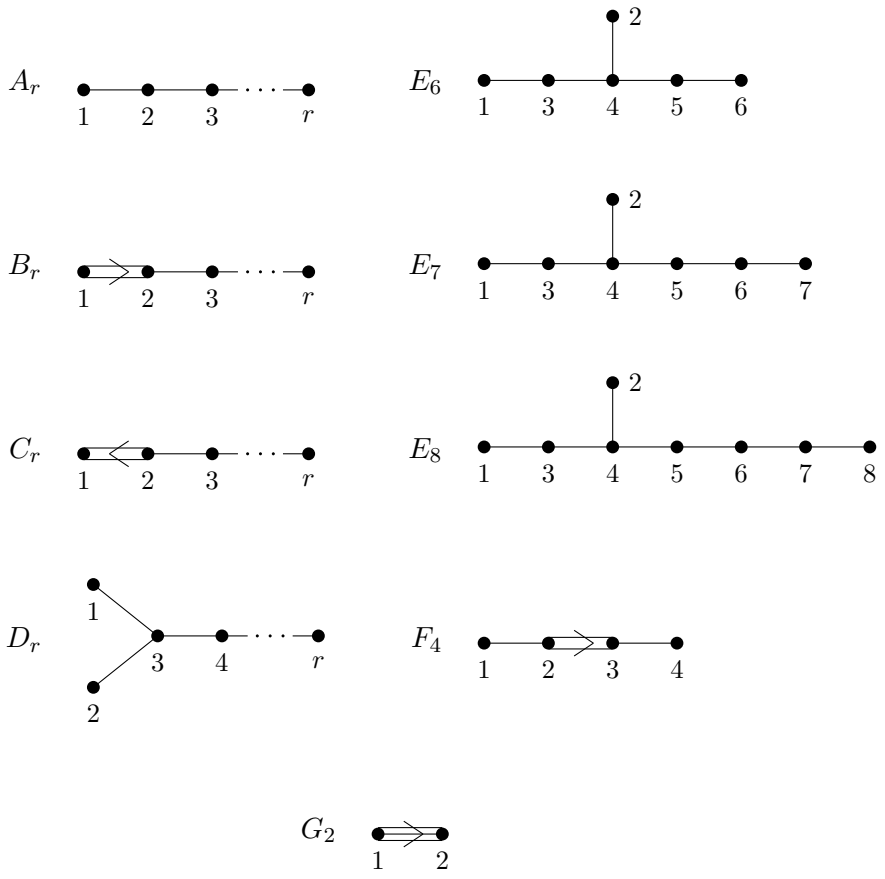


TABLE 4.2: Numbering of the simple roots

Theorem 4.11.1. *Assume $\ell \neq 2$. Write Z for the center of $G(\overline{\mathbb{F}_\ell})$ and let $\lambda = \sum_{i=1}^r a_i\omega_i$ be a q-restricted, dominant weight. Then*

- *if G is of type A_r, or D_r with odd r, or E_6, then the representations $L\left(\sum_{i=1}^r a_i\omega_i\right)$ and $L\left(\sum_{i=1}^r a_{\tau(i)}\omega_i\right)$, where the permutation τ is given by the automorphism of order two of the Dynkin diagram, are dual to each other. For any other G all representations $L(\lambda)$ are self-dual;*
- *there is an element $h \in Z$, of order at most 2, such that every self-dual module $L(\lambda)$ is symplectic if and only if h acts nontrivially on $L(\lambda)$.*

It is then relatively easy to work out which representations $L(\lambda)$ are symplectic; notice however that theorem 4.11.1 is quoted incorrectly in [65], and consequently the algorithm described in that paper to decide whether $L(\lambda)$ is symplectic or orthogonal does not yield correct results (for example, it implies the existence of symplectic representations of $\mathrm{Spin}(7, \mathbb{F}_p)$ for all sufficiently large primes p, which is not the case). The following result can be deduced directly from theorem 4.11.1, but follows more easily from an inspection of the proof of [133, Proposition 5.3]:

Corollary 4.11.2. *Assume $\ell \neq 2, 3$. In the situation of the previous theorem, the representation $L(\lambda)$ of the finite group of Lie type G^F is symplectic if and only if:*

- *G is of type A_r, $r \equiv 1 \pmod 4$, $a_i = a_{r+1-i}$ for $i = 1, \ldots, r$, and $a_{(r+1)/2}$ is odd, or*

- *G is of type B_r, $r \equiv 1, 2 \pmod 4$, and a_1 is odd, or*
- *G is of type C_r, and $a_r + a_{r-2} + \ldots + a_{r \bmod 2}$ is odd, or*
- *G is of type D_r, $r \equiv 2 \pmod 4$, $a_1 + a_2$ is odd, or*
- *G is of type E_7, and $a_2 + a_5 + a_7$ is odd.*

Proposition 4.11.3. *Let $n \leq 100$ be odd, $\ell \geq 17$ and H be a class-$\mathcal{S}$ subgroup of $\mathrm{GSp}_{2n}(\mathbb{F}_\ell)$ such that the socle of $\mathbb{P}H$ is simple of Lie type in characteristic ℓ. Then one of the following is true:*

- *up to conjugation, $soc(\mathbb{P}H)$ is the image of the $(2n-1)$-th symmetric power of the standard projective representation of $\mathrm{PSL}_2(\mathbb{F}_\ell)$;*
- *n is one of $7, 55, 63$.*

In particular, if $g \leq 100$ is an odd integer, $g \neq 7, 55, 63$, then g satisfies assumption $(*)$.

Proof. Let $\tilde{G} = \mathrm{soc}(\mathbb{P}H)$ and G^F, G, q be the associated algebraic data as in section 4.10.2. If G is of rank 1 (hence of type A_1), then $\tilde{G}$ is necessarily of the form $\mathrm{PSL}_2(q)$ for a certain $q = \ell^e$, and the case $e > 1$ is excluded by corollary 4.10.12. The same corollary also implies that in any case we have $q = \ell$, hence we can assume that G is of rank at least 2. Since we know that $\mathbb{P}H$ acts (absolutely) irreducibly on $\mathbb{P}_{2n}(\mathbb{F}_\ell)$, we are just interested in irreducible representations of G^F, that is, representations of the form $L(\lambda)$ for a certain ℓ-restricted λ.
Thus we are looking for ℓ-restricted, symplectic modules whose dimension is even, but not divisible by 4: we shall do this by looking at the tables of [65] (to which we will refer by their number in that paper), which contain a complete list of representations of degree at most 300 defined by weights that are p-restricted for at least one prime p. By corollary 4.10.27 we see that we are only interested in Lie groups of rank $r \leq \sqrt{600} < 25$. By the previous corollary, groups of type E_6, E_8, F_4, G_2 do not admit irreducible, symplectic representations. For groups of type E_7 we look at representations of even dimension: Table 6.52 shows that the smallest degree for such a representation is 912, which certainly rules out the possibility that $\dim L(\lambda) \leq 200$. We can then focus on the infinite families $A_r - D_r$, and we write $\lambda = \sum_{i=1}^r n_i \omega_i$ for the decomposition of λ along the fundamental weights.

- Type A_r ($r \geq 2$): we just need to check those r's that are congruent to 1 modulo 4, say $r = 4k+1$, and those weights $\lambda = \sum_{i=1}^r n_i \omega_i$ that satisfy $n_i = n_{r+1-i}$ for $i = 1, \ldots, r$ and $n_{2k+1} \equiv 1 \pmod 2$. Further restricting our attention to modules of dimension $\not\equiv 0 \pmod 4$, it is easy to see directly from tables 6.6-6.21 that there are no such representations with $2 \leq r \leq 17$.

  For $r \geq 21$, Table 2 shows that there are no symplectic representations of dimension at most 200; indeed this table lists all irreducible representations of dimension up to $r^3/8 > 200$, and none of them meets the requirement that $n_{(r+1)/2}$ is odd.

- Type B_r: since we are looking for symplectic representations, by the previous corollary we must have n_1 odd. Taking into account the fact that the we only need consider modules whose dimension is $\not\equiv 0 \pmod 4$, it is easy to see that no such representation exists for $r \leq 11$. Moreover, Table 2 shows that when $r \geq 12$ there are no symplectic representations of groups of type B_r of degree not exceeding $r^3 > 200$.

- Type C_r: the condition on the defining weight is now that $n_r + n_{r-2} + \dots$ be odd. We find the family of (defining) representations with highest weight $(0, 0, \dots, 1)$: these are of no interest to us, since clearly the defining representation does not give rise to a group of class $\mathcal{S}$. Apart from these, we find symplectic representations of groups of type C_3 in dimension 14 and $126 = 2 \cdot 63$, and of groups of type C_5 in dimension $110 = 2 \cdot 55$. For $r \geq 12$, Table 2 shows that (apart from the trivial and defining representations) the smallest possible degree of a nontrivial irreducible representation is $2r^2 - r - 2 > 200$.
- Type D_r: we need $r \equiv 2 \pmod 4$ and $n_1 + n_2 \equiv 1 \pmod 2$. No such representation (meeting the conditions on the dimension) exists for $r \leq 11$, and for $r \geq 12$ we see from Table 2 that the smallest possible degree of an irreducible (nontrivial, symplectic) representation is $2r^2 - r > 200$.

□

4.12 A numerical example

In this short section we consider an explicit three-dimensional Jacobian and compute a bound on the largest prime for which G_ℓ can differ from $\mathrm{GSp}_6(\mathbb{F}_\ell)$. Zywina [157] has recently given an example of a three-dimensional Jacobian having maximal (adelic) Galois action, his approach consisting essentially in making effective a previous paper by Hall [34]. Effective results based on Hall's techniques have also been obtained in [3], where an algorithm is given to test whether $G_\ell = \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ for a given abelian variety and a fixed prime ℓ. We recall that an abelian variety A/K satisfies Hall's condition if for some finite extension L of K and for some finite place v of L the fiber at v of the Néron model of $A/\mathcal{O}_L$ is semistable with toric-dimension equal to 1. Our example is fabricated precisely so as *not* to satisfy this condition, and is therefore – to the author's knowledge – the first abelian threefold not of Hall type for which the equality $G_\ell = \mathrm{GSp}_6(\mathbb{F}_\ell)$ is established for all primes larger than an explicit (albeit enormous) bound.

We now turn to the example itself. We take as abelian variety the Jacobian A of a genus 3 hyperelliptic curve C over $\mathbb{Q}$, given in an affine patch by the equation $y^2 = g(x)$, with

$$g(x) = x^7 - x^6 - 5x^5 + 4x^4 + 5x^3 - x^2 - 5x + 3.$$

The polynomial $g(x)$ has been found by referring to [50]. We shall prove that A has potentially good reduction everywhere except at $q = 45427$, and that the reduction of $A/\mathbb{Q}$ at q is semistable of toric dimension 2. Let us start by remarking that the discriminant of $g(x)$ is q^2, so C is smooth (and A has good reduction) away from 2 and q. To study the exceptional places 2 and q we shall employ the *intersection graph* of a semistable model of C:

Definition 4.12.1. Let X be a semistable curve over an algebraically closed field K. The intersection graph $\Gamma(X)$ is the (multi)graph whose vertices are the irreducible components X_i of X and whose edges are the singular points of X/K: more precisely, a singular point $x \in X$ lying on X_i and X_j defines an edge between X_i and X_j (the case $i = j$ is allowed).

Theorem 4.12.2. *([15, §9.2, Example 8]) Let X be a semistable curve over a field K. The semi-abelian variety $\mathrm{Pic}^0_{X/K}$ has toric dimension equal to* $\operatorname{rank} H^1\left(\Gamma(X_{\overline{K}}), \mathbb{Z}\right)$.

Notice now that

$$g(x) = (x + 10504)^2(x + 13963)^2 \left(x^3 + 41919x^2 + 27613x + 35727\right) \text{ in } \mathbb{F}_q[x],$$

so the reduction of C at q is semistable of toric dimension 2: indeed, our model has only ordinary double points as singularities, so the reduction is already semistable over $\mathbb{Q}_q$. Moreover, the curve is irreducible over $\overline{\mathbb{F}_q}$ and admits exactly two singular points, so the intersection graph is topologically the wedge of two copies of S^1, which shows that the toric dimension of the fiber at q is $\operatorname{rank} H^1\left(S^1 \vee S^1, \mathbb{Z}\right) = 2$. To study the reduction at 2 we shall need the following additional result:

Theorem 4.12.3. *([75, Lemma 3.2.1] and [106, Théorème 1'])* *Let K be a p-adic field with ring of integers R and denote v_p the corresponding p-adic valuation, extended to all of $\overline{K}$. Let X be the superelliptic curve given in the standard affine patch by the equation $y^p = \prod_{1\leq i\leq m}(x - x_i)$, where every x_i is in R and $(m, p) = 1$. Suppose furthermore that $v_p(x_i) = v_p(x_i - x_j) = 0$ for every pair $i \neq j$. The intersection graph of the special fiber of the stable model $\mathcal{X}$ of X is a tree.*

Take K to be the field generated over $\mathbb{Q}_2$ by the roots x_i of $g(x)$: then C/K satisfies the hypotheses of theorem 4.12.3 for $p = 2$, because $v_2\left(\prod x_i\right) = v_2(g(0)) = 0$ and

$$v_2\left(\prod_{i\neq j}(x_i - x_j)\right) = v_2(\operatorname{disc} g(x)) = 0.$$

Since trees have trivial H^1, applying theorem 4.12.2 we see that $\operatorname{Jac}(C/\mathbb{Q}_2)$ acquires good reduction over a finite extension of $\mathbb{Q}_2$: as claimed, A has potentially good reduction at 2. It follows in particular that A does not satisfy Hall's condition (over $\mathbb{Q}$, nor over any number field).
Next we check that the Galois group of $g(x)$ is the full alternating group A_7, so by [153, Theorem 2.1] we have $\operatorname{End}_{\overline{K}}(A) = \mathbb{Z}$. We then compute with Magma [16] that the characteristic polynomial of the Frobenius at 3 is $f_3(x) = 27 + 9x^5 + 6x^2 + 2x^3 + 2x^4 + x^5 + x^6$, which has Galois group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^3 \rtimes S_3$. It is interesting to observe that the characteristic polynomial of Fr_p has Galois group $(\mathbb{Z}/2\mathbb{Z})^3 \rtimes S_3$ at least for all odd primes up to 53 with the only exception of $p = 17$: a random Frobenius usually has the largest possible Galois group, so that the corresponding place satisfies assumption (3) of theorem 4.1.3. Finally, we can use [95, Théorème 2.4] to bound the Faltings height of A: the minimal discriminant of X does not exceed the discriminant of our model (namely $2^{12}q^2$), and (in the notation of [95]) we can take $e_v = 0$ to get an upper bound on $h_F(A)$. Taking into account the normalization of the Faltings height used in [95] we easily find that $h_F(A)$ does not exceed $-2.511...$
We now simply apply theorem 4.1.3 to $A/\mathbb{Q}$ and to the prime $v = 3$ to deduce that $G_\ell = \mathrm{GSp}_6(\mathbb{F}_\ell)$ for all $\ell > \exp\left(3.8 \cdot 10^8\right)$ (notice that this bound is much larger than the prime of bad reduction q).

Remark 4.12.4. The method of proof of proposition 4.7.5 produces a finite list of nonzero integers among whose prime divisors we can find all primes ℓ for which G_ℓ is of tensor product type. Actually carrying out these computations for Fr_3 rules out the possibility that G_ℓ is of tensor product type for any $\ell \geq 5$, and applying the same method to Fr_5 shows that G_3 is not of tensor product type either.

Chapter 5

The CM case

5.1 Introduction and statement of the result

The aim of this work is to study the division fields of simple abelian varieties of CM type. Recall that an abelian variety A, of dimension g and defined over a number field K, is said to admit (potential) complex multiplication, or CM for short, if there is an embedding $E \hookrightarrow \mathrm{End}_{\overline{K}}(A) \otimes \mathbb{Q}$, where E is an étale $\mathbb{Q}$-algebra of degree $2g$. We shall very often restrict to the situation of A admitting complex multiplication by E *over* K, by which we mean that $\mathrm{End}_K(A)$ is equal to $\mathrm{End}_{\overline{K}}(A)$, and of A being absolutely simple, or equivalently, of E being a number field (of degree $2g$ over $\mathbb{Q}$). The problem we discuss is that of estimating the degree $[K(A[\ell^n]) : K]$, where ℓ is a prime number and $K(A[\ell^n])$ is the field generated over K by the coordinates of the ℓ^n-torsion points of A in $\overline{K}$. As we shall see shortly, this is really a problem in the theory of Galois representations, and the seminal contributions of Shimura–Taniyama [126] and Serre–Tate [124] provide us with powerful tools for handling these representations in the CM case. Employing such tools, Silverberg studied in [128] the extension of K generated by a single torsion point of A, while Ribet gave in [110] asymptotic (non-effective) bounds on $[K(A[\ell^n]) : K]$ as $n \to \infty$. Our first result can be seen as an explicit version of the main theorem of [110]:

Theorem 5.1.1. *Let K be a number field and A/K be an abelian variety of dimension g admitting complex multiplication over K by an order in the CM field E. Denote by μ be the number of roots of unity contained in E and by $h(K)$ the class number of K. Let r be the rank of the Mumford-Tate group of A (cf. definition 5.2.10) and $\ell > \sqrt{2 \cdot g!}$ be a prime unramified in $E \cdot K$. The following inequality holds:*

$$\frac{1}{4\mu\sqrt{g!}} \cdot \ell^{nr} \leq [K(A[\ell^n]) : K] \leq \frac{5}{2}\mu \cdot h(K) \cdot \ell^{nr}.$$

Even though theorem 5.1.1 gives a good idea of the actual order of magnitude of the degree $[K(A[\ell^n]) : K]$, we can in fact prove much more precise results that apply to all primes ℓ and which are most easily described in the language of Galois representations. Recall that for every ℓ and every n there is a natural continuous action of $\mathrm{Gal}(\overline{K}/K)$ on $A[\ell^n]$, giving rise to a representation

$$\rho_{\ell^n} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(A[\ell^n]);$$

the extension $[K(A[\ell^n]) : K]$ is Galois, and its Galois group can be identified with the image G_{ℓ^n} of ρ_{ℓ^n}. Taking the inverse limit of this system of representations gives rise to the ℓ-adic representation on the Tate module $T_\ell A$,

$$\rho_{\ell^\infty} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}(T_\ell A).$$

We denote by G_{ℓ^∞} the image of ρ_{ℓ^∞} and remark that, for every n, the group G_{ℓ^n} is clearly isomorphic to the image of G_{ℓ^∞} through the canonical projection

$$\mathrm{Aut}(T_\ell A) \to \mathrm{Aut}\left(\frac{T_\ell A}{\ell^n T_\ell A}\right) \cong \mathrm{Aut}(A[\ell^n]);$$

for simplicity of exposition, we fix once and for all a $\mathbb{Z}_\ell$-basis of $T_\ell A$ and consider G_{ℓ^∞} (resp. G_{ℓ^n}) as a subgroup of $\mathrm{GL}_{2g}(\mathbb{Z}_\ell)$ (resp. of $\mathrm{GL}_{2g}(\mathbb{Z}/\ell^n\mathbb{Z})$).

We have thus reduced the problem of giving bounds on $[K(A[\ell^n]) : K]$ to that of describing G_{ℓ^n}: in trying to do so, it is natural to compare G_{ℓ^∞} with $\mathrm{MT}(A)$, the Mumford-Tate group of A (cf. definition 5.2.10). By construction, $\mathrm{MT}(A)$ is an algebraic subtorus of GL_{2g} which is only defined over $\mathbb{Q}$, so there is no obvious good definition for the group of its $\mathbb{Z}_\ell$-valued points. However, Ono [94] has shown that there is in fact a good notion of $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ (cf. definition 5.2.3), and the Mumford-Tate conjecture [84, §4] – which is a theorem for CM abelian varieties ([101] and [126]) – can be expressed by saying that, possibly after replacing K by a finite extension, G_{ℓ^∞} is a finite-index subgroup of $\mathrm{MT}(A)(\mathbb{Z}_\ell)$. For the sake of simplicity, assume for now that no extension of the base field K is necessary to attain the condition $G_{\ell^\infty} \subseteq \mathrm{MT}(A)(\mathbb{Z}_\ell)$ (our results do not depend on this assumption). The problem of estimating the degree $[K(A[\ell^n]) : K]$ is then reduced to the study of two separate quantities: the order of the finite group $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ and the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty}]$.

We treat the first problem in two important situations: when ℓ is unramified in E (a rather simple case, covered by lemma 5.2.5), and when the CM type of A is nondegenerate (theorem 5.6.1). Our result can be stated as follows:

Theorem 5.1.2. *Let A/K be an absolutely simple abelian variety of dimension g, admitting (potential) complex multiplication by the CM field E. Denote by $\mathrm{MT}(A)$ the Mumford-Tate group of A and let r be its rank.*

1. *If ℓ is unramified in E the following inequalities hold:*
$$(1 - 1/\ell)^r \ell^{nr} \leq |\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \leq (1 + 1/\ell)^r \ell^{nr}.$$
2. *Suppose $r = g + 1$. For all primes $\ell \neq 2$ and all $n \geq 1$ we have*
$$(1 - 1/\ell)^{g+1} \cdot \ell^{(g+1)n} \leq |\,\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \leq 2^g\,(1 + 1/\ell)^{g-1}\,\ell^{(g+1)n},$$
while for $\ell = 2$ and all $n \geq 1$ we have
$$\frac{1}{2^{2g+3}} \cdot 2^{(g+1)n} \leq |\,\mathrm{MT}(A)(\mathbb{Z}/2^n\mathbb{Z})| \leq \frac{1}{2}4^g \cdot 2^{(g+1)n}.$$

As for the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty}]$, our main result is as follows (cf. definition 5.2.9 for the notion of reflex norm):

Theorem 5.1.3. *(Theorem 5.5.5) Let A/K be an absolutely simple abelian variety of dimension g admitting complex multiplication over K by the CM type (E, S), and let ℓ be a prime number. If A*

has bad reduction at a place of K dividing ℓ let $\mu^* = |\mu(E)|$, the number of roots of unity in E; if on the contrary A has good reduction at all places of K of characteristic ℓ set $\mu^* = 1$. Denote by r the rank of $\mathrm{MT}(A)$ and by F the group of connected components of the kernel of the reflex norm $T_{E^*} \to T_E$, where E^* is the reflex field of E. Then:

(1) The index $[G_{\ell^\infty} : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ does not exceed $|\mu(E)| \cdot h(K)$, where $h(K)$ is the class number of K.

(2) We have $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)] \leq \mu^* \cdot [K : E^*] \cdot |F|^{2r}$.

(3) If ℓ is unramified in E and does not divide $|F|$, then the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ divides $\mu^* \cdot [K : E^*] \cdot |F|$. If ℓ is also unramified in K, the bound can be improved to $\mu^* \cdot |F|$.

Finally we have $r \leq g + 1$ and $|F| \leq f(r) \leq f(g + 1)$, where

$$f(x) = \left\lfloor 2 \left(\frac{x+1}{4} \right)^{(x+1)/2} \right\rfloor .$$

Remark 5.1.4. A few comments are in order:

- Theorem 5.1.1 follows immediately upon combining theorems 5.1.2 and 5.1.3.
- The assumption that the action of E is defined over K implies that the reflex field E^* is contained in K, see [55, Chap. 3, Theorem 1.1]. In particular, the degree $[K : E^*]$ makes sense.
- The condition $\ell \nmid |F|$ is certainly satisfied if $\ell > |F|$: in particular, it is true for all primes $\ell > f(r)$.
- Since $|F|$ is bounded by $f(g + 1)$, the degree $[K : E^*]$ does not exceed $[K : \mathbb{Q}]$, and μ^* can be controlled in terms of g alone (a trivial bound is for example $\mu^* \leq 16g^2$), we see that part (2) of theorem 5.1.3 gives a universal bound on $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ that only depends on g and $[K : \mathbb{Q}]$.
- For small values of g the function $f(g + 1)$ takes reasonably small values: we have $f(3) = 2$, $f(4) = 3$, $f(5) = 6$, $f(6) = 14$ and $f(7) = 32$.

In the special case of elliptic curves the Mumford-Tate group admits a particularly simple description, which leads to a very precise characterization of the corresponding Galois representation. Such a description can already be found (in a non-effective form) in [116, Corollaire on p.302], and the following result makes it completely explicit:

Theorem 5.1.5. *(Theorem 5.6.6) Let A/K be an elliptic curve such that $\mathrm{End}_{\overline{K}}(A)$ is an order in the imaginary quadratic field E. Denote by $\rho_\infty : \mathrm{Gal}(\overline{K}/K) \to \prod_\ell \mathrm{Aut}\, T_\ell A$ the natural adelic representation attached to A, and let G_∞ be its image. For every prime ℓ denote by C_ℓ the group $(\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times$, considered as a subgroup of $\mathrm{Aut}_{\mathbb{Z}_\ell}(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \cong \mathrm{GL}_2(\mathbb{Z}_\ell) \cong \mathrm{Aut}\, T_\ell A$, and let $N(C_\ell)$ be the normalizer of C_ℓ in $\mathrm{GL}_2(\mathbb{Z}_\ell)$.*

1. *Suppose that $E \subseteq K$: then G_∞ is contained in $\prod_\ell C_\ell$, and the index $[\prod_\ell C_\ell : G_\infty]$ does not exceed $3[K : \mathbb{Q}]$. Moreover, the equality $G_{\ell^\infty} = C_\ell$ holds for every prime ℓ unramified in K and such that A has good reduction at all places of K of characteristic ℓ.*

2. *Suppose that $E \not\subseteq K$: then G_∞ is contained in $\prod_\ell N(C_\ell)$ but not in $\prod_\ell C_\ell$, and the index $[\prod_\ell N(C_\ell) : G_\infty]$ is not finite. The intersection $H_\infty = G_\infty \cap \prod_\ell C_\ell$ has index 2 in G_∞, and the index $[\prod_\ell C_\ell : H_\infty]$ does not exceed $6[K : \mathbb{Q}]$. Moreover, the equality $G_{\ell^\infty} = N(C_\ell)$ holds for every prime ℓ unramified in $K \cdot E$ and such that A has good reduction at all places of K of characteristic ℓ.*

Finally, the constants 3 and 6 appearing in parts (1) and (2) respectively can be replaced by 1 and 2 if we further assume that the j-invariant of A is neither 0 nor 1728.

As a by-product of the proof of theorem 5.1.3 we also obtain the following proposition, which slightly strengthens a result first proved by Banaszak, Gajda and Krasoń ([7, Theorem A]) by removing both the assumption that the CM type of A is nondegenerate and the hypothesis that ℓ is completely split in K.

Proposition 5.1.6. *(Proposition 5.5.6) Let A/K be an absolutely simple abelian variety admitting complex multiplication (over K) by the CM field E, and let ℓ be a prime unramified in E. Let E^* be the reflex field of E and suppose that A has good reduction at all places of K of characteristic ℓ.*

- *The index $[\mathrm{MT}(A)(\mathbb{F}_\ell) : G_\ell \cap \mathrm{MT}(A)(\mathbb{F}_\ell)]$ divides $[K : E^*] \cdot |F|$.*
- *If ℓ is also unramified in K, then $[\mathrm{MT}(A)(\mathbb{F}_\ell) : G_\ell \cap \mathrm{MT}(A)(\mathbb{F}_\ell)]$ divides $|F|$.*

Let us conclude this introduction by giving a brief overview of the material in this chapter.
In section 5.2 we recall some fundamental notions about algebraic tori over $\mathbb{Q}$ and their $\mathbb{Z}_\ell$-points; this part also includes a brief account of the theory of abelian varieties of CM type and of their Mumford-Tate groups. In section 5.3 we apply cohomological machinery to study the map induced on $\mathbb{Z}_\ell$-points by algebraic maps between $\mathbb{Q}$-tori with good reduction at ℓ. With more effort, the method could also give results in the bad reduction setting, but the argument would become quite cumbersome and the result would not be very satisfactory for our purposes. To remedy this situation, in section 5.4 we treat the case of arbitrary reduction through a purely geometric argument inspired by [134]; it should be pointed out, however, that – in the good reduction setting – the cohomological approach gives much sharper bounds. In section 5.5 we recall a form of the Fundamental Theorem of Complex Multiplication, which gives a complete description of the Galois representations attached to A, and apply it to deduce theorem 5.1.3. In section 5.6 we give bounds on the order of $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ under the assumption that A is of nondegenerate type, i.e. that $\operatorname{rank} \mathrm{MT}(A) = \dim A + 1$. Finally, in the short section 5.7 we give a simple example that shows that the optimal bound on $\ell^{n \operatorname{rank} \mathrm{MT}(A)} \,/\, [K(A[\ell^n]) : K]$ grows at least exponentially fast in g, so that our bounds are not too far from the truth.

5.2 Preliminaries on algebraic tori

Recall that over a perfect field k there is an equivalence of categories between algebraic tori and finitely generated, torsion-free, continuous $\mathrm{Gal}(\overline{k}/k)$-modules: if T is a k-torus, the corresponding $\mathrm{Gal}(\overline{k}/k)$-module is the group of characters $\hat{T} = \mathrm{Hom}\left(T_{\overline{k}}, \mathbb{G}_{m,\overline{k}}\right)$. Also recall that this construction extends to an equivalence between finitely generated, continuous $\mathrm{Gal}(\overline{k}/k)$-modules and k-group schemes of multiplicative type; we will make use of this fact to study the kernel of the reflex norm. We now introduce a family of $\mathbb{Q}$-algebraic tori that will be especially relevant for us:

Definition 5.2.1. If E is any number field we set $T_E = \mathrm{Res}_{E/\mathbb{Q}}(\mathbb{G}_{m,\mathbb{Q}})$.

The torus T_E is of rank $[E:\mathbb{Q}]$, and it admits a very simple description in terms of characters: it is the $\mathbb{Q}$-torus that corresponds to the free module over the set $\mathrm{Hom}(E, \overline{\mathbb{Q}})$, endowed with its natural (right) $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-action.

Proposition 5.2.2. *Let E be a number field. The torus T_E has good reduction at all the primes not dividing* $\mathrm{disc}(E)$.

Proof. By the Galois criterion ([88, Proposition 1.1]), T_E has good reduction at ℓ if and only if the inertia group at (a place of $\overline{\mathbb{Q}}$ over) ℓ acts trivially on $\widehat{T_E}$. In the present case $\widehat{T_E}$ is the free module over $\mathrm{Hom}(E, \overline{\mathbb{Q}})$, so if we let L be the Galois closure of E in $\overline{\mathbb{Q}}$ the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $\widehat{T_E}$ factors through its finite quotient $\mathrm{Gal}(L/\mathbb{Q})$. Now if a prime ℓ is unramified in E it is also unramified in L, hence the inertia at ℓ has trivial image in $\mathrm{Gal}(L/\mathbb{Q})$ and T_E has good reduction at ℓ, as claimed. □

5.2.1 Points of tori with values in $\mathbb{Z}_\ell$ and $\mathbb{Z}/\ell^n\mathbb{Z}$

We briefly discuss the various possible definitions for the group of $\mathbb{Z}_\ell$-valued points of a $\mathbb{Q}_\ell$-torus; our main reference for this section is [110, §2]. Let T be a $\mathbb{Q}_\ell$-torus, not necessarily having good reduction over $\mathbb{F}_\ell$. We fix a finite Galois extension L of $\mathbb{Q}_\ell$ that splits T, and we regard $\hat{T}$ as a Γ-module, where $\Gamma := \mathrm{Gal}(L/\mathbb{Q}_\ell)$. Also notice that a character $\chi \in \hat{T}$ can in particular be considered as a homomorphism $\chi : T(L) \to L^\times$.

Definition 5.2.3. Following Ono ([94, §2]), we define $T(\mathbb{Z}_\ell)$ to be $\mathrm{Hom}_\Gamma\left(\hat{T}, \mathcal{O}_L^\times\right)$, the group of Γ-equivariant morphisms (of abelian groups) of $\hat{T}$ in $\mathcal{O}_L^\times$. Equivalently, $T(\mathbb{Z}_\ell)$ is the maximal compact subgroup of $T(\mathbb{Q}_\ell)$.

If furthermore we suppose that T has good reduction, then it is known ([138, Theorem 2 on p.109]) that there exists a $\mathbb{Z}_\ell$-model $\mathcal{T}$ of T (that is, a commutative smooth group scheme over $\mathrm{Spec}(\mathbb{Z}_\ell)$ whose generic fiber is T). As pointed out in [110, Remark 2.2], in this case the $\mathbb{Z}_\ell$-points of T in the sense of Ono agree with the $\mathbb{Z}_\ell$-valued points of $\mathcal{T}$, so that we are free to use whichever definition we find more convenient. When a smooth model $\mathcal{T}$ exists we can also give the following definition:

Definition 5.2.4. If T has good reduction, the $\mathbb{Z}/\ell^n\mathbb{Z}$-points of T are the $\mathbb{Z}/\ell^n\mathbb{Z}$-valued points of its smooth $\mathbb{Z}_\ell$-model $\mathcal{T}$.

We still need to discuss the meaning of $T(\mathbb{Z}/\ell^n\mathbb{Z})$ when T does *not* have good reduction. The construction in this case is again due to Ono. For $n \geq 0$, we define subgroups of $T(\mathbb{Q}_\ell)$ by the rule

$$T(1+\ell^n\mathbb{Z}_\ell) = \left\{ x \in T(\mathbb{Q}_\ell) \;\middle|\; v_\ell(\chi(x)-1) \geq n \quad \forall \chi \in \hat{T} \right\}.$$

We simply write $T(\mathbb{Z}_\ell)$ for the group corresponding to $n = 0$: it can be easily checked that this definition agrees with our previous ones. We can now set $T(\mathbb{Z}/\ell^n\mathbb{Z}) = \dfrac{T(\mathbb{Z}_\ell)}{T(1+\ell^n\mathbb{Z}_\ell)}$; once again, when T has a smooth $\mathbb{Z}_\ell$-model $\mathcal{T}$, the group $T(\mathbb{Z}/\ell^n\mathbb{Z})$ agrees with $\mathcal{T}(\mathbb{Z}/\ell^n\mathbb{Z})$. Finally, when T is a $\mathbb{Q}$-torus we define $T(\mathbb{Z}/\ell^n\mathbb{Z})$ to be the group of $\mathbb{Z}/\ell^n\mathbb{Z}$-points of $T \otimes \mathbb{Q}_\ell$. We conclude this discussion with the following well-known lemma:

Lemma 5.2.5. *Let $T/\mathbb{Q}_\ell$ have good reduction. For every positive integer n we have*

$$(1-1/\ell)^{\dim T}\ell^{n\dim T} \leq |T(\mathbb{Z}/\ell^n\mathbb{Z})| \leq (1+1/\ell)^{\dim T}\ell^{n\dim T}.$$

Proof. A combination of Hensel's lemma and [138, Theorem 2 on p.104]; for further details, we refer the reader to [39, Lemme 2.1 and Proposition 2.2]. □

5.2.2 CM types and reflex norm

We briefly recall the notions of CM type, of reflex type, and of reflex norm; we refer the reader to [110, §3] for further details. Let E be a CM field of degree $2g$ and $\tilde{E}$ be its Galois closure in $\overline{\mathbb{Q}}$, and write G, H for the Galois groups $\mathrm{Gal}(\tilde{E}/\mathbb{Q})$ and $\mathrm{Gal}(\tilde{E}/E)$ respectively. We denote by τ the complex conjugation of $\mathbb{C}$, or any of its restrictions, and we take the convention that the set $\mathrm{Hom}(E, \overline{\mathbb{Q}})$ be identified with the coset space $H\backslash G$.

Lemma 5.2.6. *The degree $[\tilde{E} : \mathbb{Q}]$ divides $2^g g!$.*

Proof. Let E_0 be the maximal totally real subfield of E and $a \in E_0$ be such that $E = E_0(\sqrt{a})$. Let $\tilde{E}_0$ be the Galois closure of E_0 and $a_1 = a, \ldots, a_k \in \tilde{E}_0$ be the conjugates of a over $\mathbb{Q}$, where $k \leq [E_0 : \mathbb{Q}] = g$. It is clear that $\tilde{E}$ is generated over $\tilde{E}_0$ by $\sqrt{a_1}, \ldots, \sqrt{a_k}$, so $[\tilde{E} : \mathbb{Q}]$ divides $[\tilde{E}_0 : \mathbb{Q}] \cdot 2^k$. As $[\tilde{E}_0 : \mathbb{Q}] \mid g!$ and $k \leq g$ the lemma follows. □

Definition 5.2.7. A CM-type for the CM field E is a subset S of $H\backslash G$ such that $S \cap \tau(S) = \emptyset$ and $H\backslash G = S \cup \tau(S)$.

Let S be a CM type for E and $\tilde{S}$ be the inverse image of S in G, i.e. $\tilde{S} = \left\{ g \in G \;\middle|\; Hg \in S \right\}$. We set $H' = \left\{ g \in G \;\middle|\; \tilde{S}g = \tilde{S} \right\}$ and let E^* be the fixed field of H'; we then set $\tilde{R} = \left\{ s^{-1} \;\middle|\; s \in \tilde{S} \right\}$ and let R be the image of $\tilde{R}$ in $H'\backslash G \cong \mathrm{Hom}\left(E^*, \overline{\mathbb{Q}}\right)$. It is not hard to check that R is a CM type for E^*.

Definition 5.2.8. The pair (E^*, R) is called the **reflex type** of (E, S).

Finally, a CM type (E, S) is called simple if the equality

$$H = \left\{ g \in G \;\middle|\; g\tilde{S} = \tilde{S} \right\}$$

holds. We are now ready to define the reflex norm:

Definition 5.2.9. Let (E, S) be a CM type, $\tilde{E}$ the Galois closure of $E/\mathbb{Q}$ and (E^*, R) the reflex type of (E, S). The **reflex norm** associated with (E, S) is the $\mathbb{Q}$-morphism

$$\Phi_{(E,S)} : T_{E^*} \to T_E$$

of algebraic tori given on characters by

$$\begin{array}{rccc} \Phi^*_{(E,S)} & : & \widehat{T_E} & \to & \widehat{T^*_E} \\ & & [g] & \mapsto & \sum_{r\in R}[rg], \end{array}$$

where $[g]$ (resp. $[rg]$) is the embedding of E (resp. E^*) in $\overline{\mathbb{Q}}$ induced by the automorphism $g \in \mathrm{Gal}(\tilde{E}/\mathbb{Q})$ (resp. $rg \in \mathrm{Gal}(\tilde{E}/\mathbb{Q})$).

5.2.3 The Mumford-Tate group

Our interest in the reflex norm stems from the fact that it allows us to define the Mumford-Tate group of a CM abelian variety rather directly. Before doing so, however, we need to recall how one associates a CM type with a CM abelian variety.

Let A/K be an absolutely simple abelian variety, admitting complex multiplication (over $\overline{K}$) by the field E. The tangent space at the identity of $A_{\overline{K}}$ is a $\overline{K}$-module and an E-module, and the two actions are compatible: it follows that this tangent space is a $(E \otimes \overline{K})$-bimodule, so it decomposes as $T_{\mathrm{id}}A_{\overline{K}} \cong \prod_{\varphi\in S} \overline{K}_\varphi$, where $\overline{K}_\varphi$ is a 1-dimensional $\overline{K}$-vector space on which E acts through the embedding $\varphi : E \hookrightarrow \overline{K}$. The set S of embeddings that appear in this decomposition can be shown to be a CM type for E, and in this case we say that A admits complex multiplication by the CM type (E, S). When furthermore we have $\mathrm{End}_K(A) = \mathrm{End}_{\overline{K}}(A)$ we say that A admits complex multiplication by (E, S) *over* K.

Definition 5.2.10. Let A/K be an absolutely simple abelian variety admitting complex multiplication (over $\overline{K}$) by the CM type (E, S), and let (E^*, R) be the reflex type. We define the Mumford-Tate torus $\mathrm{MT}(A)$ to be the image of the reflex norm $\Phi_{(E,S)} : T_{E^*} \to T_E$.

Remark 5.2.11. The Mumford-Tate group of A is in fact a purely geometric object – it can described in terms of the Hodge structure associated with the complex abelian variety $A_{\mathbb{C}}$. In particular, it is insensitive to extensions of the base field K.

Remark 5.2.12. It is known that the rank of $\mathrm{MT}(A)$ is at most $g + 1$. When equality holds, the CM type is said to be nondegenerate, and the Mumford-Tate group has a very simple description in terms of E: if τ denotes complex conjugation on E, for any $\mathbb{Q}$-algebra B the B-points of $\mathrm{MT}(A)$ are given by

$$\mathrm{MT}(A)(B) = \left\{ x \in (E \otimes_{\mathbb{Q}} B)^\times \;\middle|\; x\tau(x) \in B^\times \right\}.$$

For all these facts see for example [110], Proposition 3.3 and the remarks following it.

5.2.4 The group of connected components of $\ker \Phi_{(E,S)}$

An object which will be crucial to our study is the kernel of the reflex norm $\Phi_{(E,S)}$: in this short subsection we establish a bound on the order of its group of components. The bound is ultimately a consequence of Hadamard's inequality, which is the main tool used to establish the following lemma:

Lemma 5.2.13. *Let A be a $n \times n$ integral matrix all of whose entries are in $\{0,1\}$. The following inequality holds:*

$$|\det A| \leq \lfloor 2^{-n}(n+1)^{(n+1)/2} \rfloor.$$

Proof. Consider the matrix

$$B(A) = \left(\begin{array}{c|c} 1 & 1 \cdots 1 \\ \hline 0 \\ \vdots & 2A \\ 0 \end{array}\right).$$

It is clear by definition that $\det B(A) = 2^n \det(A)$. Consider the matrix $H(A)$ obtained from $B(A)$ by subtracting the first row to each of the others. Clearly $H(A)$ and $B(A)$ have the same determinant, and furthermore all the entries of $H(A)$ are in $\{\pm 1\}$. In particular, the L^2-norm of every row of $H(A)$ is $\sqrt{n+1}$, so Hadamard's inequality implies

$$|\det A| = 2^{-n} |\det B(A)| = 2^{-n} |\det H(A)| \leq 2^{-n}(n+1)^{(n+1)/2}.$$

The claim then follows from the fact that $\det(A)$ is an integer. □

Lemma 5.2.14. *Let $T : \mathbb{Z}^n \to \mathbb{Z}^m$ be a linear map, represented in the standard bases by a matrix A all of whose entries are in $\{0,1\}$. Let Y be the image of T, denote by k the rank of Y, and let Z be given by $Z = \left\{ z \in \mathbb{Z}^m \;\middle|\; \exists q \in \mathbb{Z} \text{ such that } qz \text{ belongs to } Y \right\}$. The quotient Z/Y, which is isomorphic to the torsion part of $\mathbb{Z}^m/Y$, has order at most $\lfloor 2^{-k}(k+1)^{(k+1)/2} \rfloor$.*

Proof. The order of Z/Y is given by $\gcd\left\{\det(A_k) \;\middle|\; A_k \text{ is a minor of } A \text{ of size } k\right\}$. Lemma 5.2.13 ensures that the determinant of every minor of size k does not exceed $\lfloor 2^{-k}(k+1)^{(k+1)/2} \rfloor$, and the lemma follows. □

Proposition 5.2.15. *Let C be the group of multiplicative type defined by the exact sequence*

$$1 \to C \to T_{E^*} \xrightarrow{\Phi_{(E,S)}} \mathrm{MT}(A) \to 1$$

and let $\hat{C}$ be its character group. Suppose $\mathrm{MT}(A)$ has rank r. The torsion subgroup of $\hat{C}$ has order at most $\lfloor 2^{-r}(r+1)^{(r+1)/2} \rfloor$.

Proof. Let Y be the image of $\Phi^*_{(E,S)} : \hat{T}_E \to \hat{T}_{E^*}$ and

$$Z = \left\{ \chi \in \hat{T}_{E^*} \;\middle|\; \exists n \in \mathbb{Z} \text{ such that } n\chi \in Y \right\}.$$

The torsion subgroup of $\hat{C}$ is isomorphic to Z/Y. Moreover, it is apparent from definition 5.2.9 that the matrix representing $\Phi^*_{(E,S)}$ in the natural bases of $\widehat{T_{E^*}}, \widehat{T_E}$ has entries in $\{0,1\}$, so the proposition follows from lemma 5.2.14. □

5.3 Cohomology and integral points of tori

The purpose of this section is to study the map induced on $\mathbb{Z}_\ell$-points by a surjection of tori over $\mathbb{Q}_\ell$. More precisely, we let $T \xrightarrow{\beta} T'' \to 1$ be a surjection of $\mathbb{Q}_\ell$-algebraic tori, and we assume that T has good reduction. We let T' be the kernel of β, which is in general just a group of multiplicative type (and not necessarily a torus), and write F for the torsion subgroup of its character group $\hat{T}'$.

We also denote by a the rank of T', so that we have an isomorphism of abelian groups $\hat{T}'/F \cong \mathbb{Z}^a$. Finally, we fix a finite *unramified* Galois extension L of $\mathbb{Q}_\ell$ that splits T, and we let Γ denote the Galois group of L over $\mathbb{Q}_\ell$. It is also useful to introduce the following notation:

Notation. If n is any integer and ℓ is a prime we write $|n|_\ell$ for $\ell^{-v_\ell(n)}$. When M is a finite group we also write $|M|_\ell$ for $\ell^{-v_\ell(|M|)}$.

With this notation we shall show:

Proposition 5.3.1. *The cokernel of $T(\mathbb{Z}_\ell) \xrightarrow{\beta} T''(\mathbb{Z}_\ell)$ has order dividing $|F| \cdot |F|_\ell^{-[L:\mathbb{Q}_\ell]}$.*

The proof is given below in §5.3.2, and relies mainly on the basic tools of Galois cohomology, together with the following classical theorem of Nakayama (cf. for example [125, §2, Theorem 32]):

Theorem 5.3.2. *Let A and B be modules over the finite group G. Assume that A is cohomologically trivial. In order for $\mathrm{Hom}(B, A)$ to be cohomologically trivial it is necessary and sufficient that $\mathrm{Ext}^1(B, A)$ be cohomologically trivial. In particular, if B is $\mathbb{Z}$-free, then $\mathrm{Hom}(B, A)$ is cohomologically trivial.*

5.3.1 Preliminaries on p-adic fields

The following two lemmas are certainly well-known, but for lack of an easily accessible reference we prefer to include a short proof.

Lemma 5.3.3. *Let L be a finite extension of $\mathbb{Q}_\ell$ with ring of integers $\mathcal{O}_L$, and let n be a positive integer. The quotient $\mathcal{O}_L/\mathcal{O}_L^{\times n}$ has order dividing $n \cdot |n|_\ell^{-[L:\mathbb{Q}_\ell]}$.*

Proof. We regard all the involved groups as $\mathbb{Z}/n\mathbb{Z}$-modules with trivial action, and denote by h_n the associated Herbrand quotient, that is to say for every finite $\mathbb{Z}/n\mathbb{Z}$-module M we set

$$h_n(M) := \frac{|\hat{H}^0(\mathbb{Z}/n\mathbb{Z}, M)|}{|H^1(\mathbb{Z}/n\mathbb{Z}, M)|}.$$

As $\mathcal{O}_L^1$, the subgroup of principal units of $\mathcal{O}_L$, has finite index in $\mathcal{O}_L^\times$ (and the Herbrand quotient is invariant by passage to finite-index subgroups), we have $h_n(\mathcal{O}_L^\times) = h_n(\mathcal{O}_L^1)$. On the other hand, $\mathcal{O}_L^1$ contains a subgroup of finite index that is isomorphic to $\mathcal{O}_L$ ([115, Chapitre XIV, prop. 10]), so $h_n(\mathcal{O}_L^\times) = h_n(\mathcal{O}_L^1) = h_n(\mathcal{O}_L)$. Furthermore, $H^1(\mathbb{Z}/n\mathbb{Z}, \mathcal{O}_L^\times) = \mathrm{Hom}\left(\mathbb{Z}/n\mathbb{Z}, \mathcal{O}_L^\times\right) = \mathcal{O}_L^\times[n]$ has order dividing n, while $H^1(\mathbb{Z}/n\mathbb{Z}, \mathcal{O}_L) = \mathcal{O}_L[n] = 0$. The lemma then follows easily because the quantity $\left|\frac{\mathcal{O}_L^\times}{\mathcal{O}_L^{\times n}}\right| = h^1(\mathbb{Z}/n\mathbb{Z}, \mathcal{O}_L^\times) \cdot h_n\left(\mathcal{O}_L^\times\right)$ divides

$$n \cdot h_n(\mathcal{O}_L) = n\frac{|\mathcal{O}_L/n\mathcal{O}_L|}{h^1(\mathbb{Z}/n\mathbb{Z}, \mathcal{O}_L)} = n \cdot |n|_\ell^{-[L:\mathbb{Q}_\ell]}.$$

□

Lemma 5.3.4. *Let F be a finite abelian group and L be a finite extension of $\mathbb{Q}_\ell$. Then $|\mathrm{Ext}^1(F, \mathcal{O}_L^\times)|$ divides $|F| \cdot |F|_\ell^{-[L:\mathbb{Q}_\ell]}$.*

Proof. Writing F as $\bigoplus_i \frac{\mathbb{Z}}{d_i\mathbb{Z}}$ we have

$$\mathrm{Ext}^1\left(F, \mathcal{O}_L^\times\right) \cong \prod_i \mathrm{Ext}^1\left(\frac{\mathbb{Z}}{d_i\mathbb{Z}}, \mathcal{O}_L^\times\right) \cong \prod_i \frac{\mathcal{O}_L^\times}{\mathcal{O}_L^{\times d_i}}.$$

The result follows from the previous lemma. □

5.3.2 Proof of proposition 5.3.1

Note that – since $L/\mathbb{Q}_\ell$ is unramified – the group $\mathcal{O}_L^\times$ is a cohomologically trivial Γ-module (cf. for example [90, Prop. 7.1.2 (ii)]). As $\hat{T}$ and $\hat{T}''$ are free abelian groups, Nakayama's theorem implies in particular that $\mathrm{Hom}\left(\hat{T}, \mathcal{O}_L^\times\right)$ and $\mathrm{Hom}\left(\hat{T}'', \mathcal{O}_L^\times\right)$ are cohomologically trivial Γ-modules. We will make extensive use of this fact. The character groups of T, T', T'' fit into an exact sequence

$$0 \to \hat{T}'' \to \hat{T} \to \hat{T}' \to 0;$$

applying the functor $\mathrm{Hom}\left(-, \mathcal{O}_L^\times\right)$ gives another exact sequence

$$0 \to \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right) \to \mathrm{Hom}\left(\hat{T}, \mathcal{O}_L^\times\right) \to \mathrm{Hom}\left(\hat{T}'', \mathcal{O}_L^\times\right) \to \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right) \to 0,$$

where the following Ext term vanishes since $\hat{T}$ is free. If we let

$$I := \mathrm{Image}\left(\mathrm{Hom}\left(\hat{T}, \mathcal{O}_L^\times\right) \to \mathrm{Hom}\left(\hat{T}'', \mathcal{O}_L^\times\right)\right),$$

the previous sequence gives rise to the two exact sequences

$$0 \to \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right) \to \mathrm{Hom}\left(\hat{T}, \mathcal{O}_L^\times\right) \to I \to 0 \tag{5.1}$$

and

$$0 \to I \to \mathrm{Hom}\left(\hat{T}'', \mathcal{O}_L^\times\right) \to \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right) \to 0. \tag{5.2}$$

Writing as usual I^Γ for $H^0(\Gamma, I)$, the long exact sequences in Galois cohomology associated with equations (5.1) and (5.2) give

$$0 \to \mathrm{Hom}_\Gamma\left(\hat{T}', \mathcal{O}_L^\times\right) \to T(\mathbb{Z}_\ell) \to I^\Gamma \to H^1\left(\Gamma, \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right)\right) \to 0, \tag{5.3}$$

$$0 \to H^1(\Gamma, I) \to H^2\left(\Gamma, \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right)\right) \to 0, \tag{5.4}$$

and

$$0 \to I^\Gamma \to T''(\mathbb{Z}_\ell) \to H^0\left(\Gamma, \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right)\right) \to H^1(\Gamma, I) \to 0, \tag{5.5}$$

where we have used the fact that $\mathrm{Hom}\left(\hat{T}, \mathcal{O}_L^\times\right)$ and $\mathrm{Hom}\left(\hat{T}'', \mathcal{O}_L^\times\right)$ are cohomologically trivial. Also notice that we have an exact sequence of Γ-modules

$$0 \to F \to \hat{T}' \to \hat{T}'/F \to 0 \tag{5.6}$$

where $\hat{T}'/F \cong \mathbb{Z}^a$ is free. We can then apply $\mathrm{Hom}\left(-, \mathcal{O}_L^\times\right)$ to (5.6) to get

$$0 \to \mathrm{Hom}\left(\hat{T}'/F, \mathcal{O}_L^\times\right) \to \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right) \to \mathrm{Hom}\left(F, \mathcal{O}_L^\times\right) \to 0,$$

and since $\mathrm{Hom}\left(\hat{T}'/F, \mathcal{O}_L^\times\right)$ is again cohomologically trivial by theorem 5.3.2 we deduce that for every $n \geq 1$ we have canonical isomorphisms

$$H^n\left(\Gamma, \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right)\right) \xrightarrow{\sim} H^n\left(\Gamma, \mathrm{Hom}\left(F, \mathcal{O}_L^\times\right)\right). \tag{5.7}$$

Straightforward manipulations of sequences (5.3) and (5.5) show that

$$\left|\mathrm{coker}\left(T(\mathbb{Z}_\ell) \to T''(\mathbb{Z}_\ell)\right)\right| = \frac{h^0\left(\Gamma, \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right)\right) \cdot h^1\left(\Gamma, \mathrm{Hom}\left(\hat{T}', \mathcal{O}_L^\times\right)\right)}{h^1(\Gamma, I)}.$$

For the sake of notational simplicity set $M = \mathrm{Hom}\left(F, \mathcal{O}_L^\times\right)$. Using (5.4) and (5.7) we arrive at

$$\left|\mathrm{coker}\left(T(\mathbb{Z}_\ell) \to T''(\mathbb{Z}_\ell)\right)\right| = \frac{h^0\left(\Gamma, \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right)\right) \cdot h^1\left(\Gamma, M\right)}{h^2(\Gamma, M)}. \tag{5.8}$$

Observe now that the group Γ is cyclic (since it is the Galois group of an unramified extension) and the module M is finite: as it is well-known, the Tate cohomology $\hat{H}^n$ of a cyclic group with values in a finite module is 2-periodic in n. Moreover, the Herbrand quotient $\frac{\left|\hat{H}^0(\Gamma,M)\right|}{\left|\hat{H}^1(\Gamma,M)\right|}$ equals 1 since M is finite, and therefore $h^2(\Gamma, M) = \left|\hat{H}^0(\Gamma, M)\right| = h^1(\Gamma, M)$ (for all these facts cf. for example [90, §I.7]). Using this equality in (5.8) we finally find $\left|\mathrm{coker}\left(T(\mathbb{Z}_\ell) \to T''(\mathbb{Z}_\ell)\right)\right| = h^0\left(\Gamma, \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right)\right)$. Proposition 5.3.1 then follows from the fact that $h^0\left(\Gamma, \mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right)\right)$ divides

$$\left|\mathrm{Ext}^1\left(\hat{T}', \mathcal{O}_L^\times\right)\right| = \left|\mathrm{Ext}^1\left(\mathbb{Z}^a \oplus F, \mathcal{O}_L^\times\right)\right| = \left|\mathrm{Ext}^1\left(F, \mathcal{O}_L^\times\right)\right|$$

and from lemma 5.3.4.

5.4 The cokernel of an isogeny, without the good reduction assumption

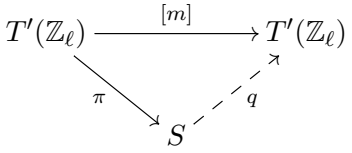
Let T, T' be $\mathbb{Q}_\ell$-tori and $\lambda : T \to T'$ be a $\mathbb{Q}_\ell$-isogeny. We do not assume that T or T' has good reduction, and for the purposes of this section we define the $\mathbb{Z}_\ell$-points of a $\mathbb{Q}_\ell$-torus to be the maximal compact subgroup of $T(\mathbb{Q}_\ell)$ (cf. definition 5.2.3). Our aim is again to bound the order of coker $\left(T(\mathbb{Z}_\ell) \xrightarrow{\lambda} T'(\mathbb{Z}_\ell)\right)$, in terms of the degree m of λ and of $\dim T = \dim T' =: d$. Cohomological tools could again be used to investigate the problem, but we find that an entirely different approach (through p-adic differential geometry) yields simpler and more effective proofs; the method is inspired by [134], see especially lemma 4.4 in *op. cit.*

Proposition 5.4.1. *Let T, T' be $\mathbb{Q}_\ell$-tori of dimension d and $\lambda : T \to T'$ be an isogeny of degree m. The order of* coker $\left(T(\mathbb{Z}_\ell) \xrightarrow{\lambda} T'(\mathbb{Z}_\ell)\right)$ *is at most $m^d \cdot |m|_\ell^{-d}$.*

Proof. Notice first that λ fits into a commutative diagram

$$\begin{array}{ccc} T'(\mathbb{Z}_\ell) & \overset{\lambda^\vee}{\dashrightarrow} & T(\mathbb{Z}_\ell) \\ & {\scriptstyle [m]} \searrow \quad \swarrow {\scriptstyle \lambda} & \\ & T'(\mathbb{Z}_\ell) & \end{array}$$

and therefore it is enough to bound the cokernel of $[m] : T'(\mathbb{Z}_\ell) \to T'(\mathbb{Z}_\ell)$. Fix now a Haar measure μ on $T'(\mathbb{Q}_\ell)$, normalized in such a way that $\mu(T'(\mathbb{Z}_\ell)) = 1$.
Consider the kernel K of $[m]$ (as a subgroup of $T'(\mathbb{Z}_\ell)$, not as a group scheme) and the quotient $S = T'(\mathbb{Z}_\ell)/K$, and note that $\pi : T'(\mathbb{Z}_\ell) \to S$ is a covering map. We denote by μ_S the measure on S given by $\mu_S(A) = \frac{1}{|K|}\mu\left(\pi^{-1}(A)\right)$: it can also be interpreted as the measure induced on S by the (Haar) volume form of $T'(\mathbb{Z}_\ell)$, which passes to the quotient since it is translation-invariant. The volume of S (for the measure μ_S) is $\frac{\mathrm{vol}(T'(\mathbb{Z}_\ell))}{|K|} = \frac{1}{|K|}$, and we have an ℓ-adic analytic map $q : S \to T'(\mathbb{Z}_\ell)$ such that the following diagram commutes:



Clearly q is an ℓ-adic analytic embedding and we have $\operatorname{Image} q = \operatorname{Image}[m] =: I$. We have the following immediate equality:

$$\operatorname{vol}(I) = \frac{1}{|T'(\mathbb{Z}_\ell)/I|} \operatorname{vol}\left(T'(\mathbb{Z}_\ell)\right) = \frac{1}{|T'(\mathbb{Z}_\ell)/I|}. \tag{5.9}$$

On the other hand, a simple computation in coordinates shows $q^*\mu = |m|_\ell^d \, \mu_S$: we can parametrize a neighbourhood of $g \in T'(\mathbb{Z}_\ell)$ by $x \mapsto g\exp(x)$ (for x varying in some small neighbourhood of 0 in the Lie algebra of $T'(\mathbb{Q}_\ell)$), and composing with π this also induces a parametrization of a neighbourhood of $\pi(g) \in S$. In these coordinates the map q is simply multiplication by m, so its Jacobian determinant is $|m|_\ell^d$ and the change of variables formula for ℓ-adic integration gives the required result. This yields

$$\operatorname{vol}(I) = \int_I d\mu = \int_{q(S)} d\mu = \int_S d(q^*\mu) = \int_S |m|_\ell^d \, d\mu_S = |m|_\ell^d \, \frac{1}{|K|},$$

and comparing this equality with equation (5.9) gives

$$\left|\operatorname{coker}\left(T'(\mathbb{Z}_\ell) \xrightarrow{[m]} T'(\mathbb{Z}_\ell)\right)\right| = |T'(\mathbb{Z}_\ell)/I| = \frac{1}{\operatorname{vol}(I)} = \frac{|K|}{|m|_\ell^d}.$$

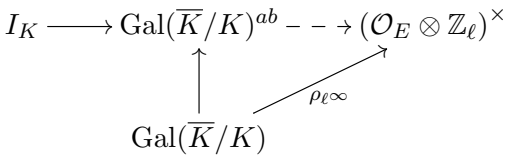
Finally, it is clear that $|K| \leq \left|T'(\overline{\mathbb{Q}_\ell})[m]\right| = m^d$, and this finishes the proof. □

5.5 Description of the Galois representation

Let A/K be an absolutely simple g-dimensional CM abelian variety admitting complex multiplication (over K) by the CM type (E, S). Let $\tilde{E}$ be the Galois closure of E, denote by (E^*, R) the reflex type of (E, S), and let ℓ be a prime number. It is known that – since the action of E is defined over K – the reflex field E^* is contained in K ([55, Chap. 3, Theorem 1.1]), and by [124, Corollary 2 to Theorem 5], the ℓ-adic Galois representation attached to A can be viewed as a map

$$\rho_{\ell\infty} : \operatorname{Gal}(\overline{K}/K) \to (\operatorname{End}_K(A) \otimes \mathbb{Z}_\ell)^\times \hookrightarrow (\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times .$$

We denote by $G_{\ell\infty}$ the image of $\rho_{\ell\infty}$. We now recall the description of $\rho_{\ell\infty}$ coming from the fundamental theorem of complex multiplication, and refer the reader to [110, §4] and [124] for further details. Let I_K be the group of idèles of K. As $(\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times$ is commutative, there is a factorization



which (by class field theory) allows us to regard $\rho_{\ell\infty}$ as a map from I_K to $(\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times$. Let us introduce some notation: we write $\mu(E)$ for the group of roots of unity in E, and if v is a place of K we write $\mathcal{O}_{K,v}$ for the completion at v of the ring of integers of K. If v is furthermore finite we

denote by p_v its residual characteristic; we also let Ω_K be the set of all finite places of K. If F is a number field we denote by F_ℓ the algebra $F \otimes \mathbb{Q}_\ell$, and for an idèle $a \in I_K$ we write a_ℓ for the component of a in $K_\ell \cong \prod_{p_v=\ell} K_v^\times$. With this notation, the map ρ_{ℓ^∞} is described very precisely by the following theorem:

Theorem 5.5.1. *([124, Theorems 6, 10 and 11]) There exists a unique continuous homomorphism $\varepsilon : I_K \to E^\times$ such that, for all finite places v of K, the group $\varepsilon\left(\mathcal{O}_{K,v}^\times\right)$ is contained in $\mu(E)$, and*

$$\rho_{\ell^\infty}(a) = \varepsilon(a)\Phi_{(E,S)}\left(N_{K_\ell/E_\ell^*}(a_\ell)\right)^{-1}$$

for all $a \in I_K$. If furthermore $v \in \Omega_K$ is a place of good reduction for A, then $\varepsilon\left(\mathcal{O}_{K,v}^\times\right)$ is trivial.

We now consider the restriction of ρ_{ℓ^∞} to $K^\times \cdot \prod_{v\in\Omega_K} \mathcal{O}_{K,v}^\times$: as it is well-known (cf. for example [89, Proposition 2.3]), this is the group of idèles of H, the Hilbert class field of K. In terms of Galois groups, this has the effect of restricting ρ_{ℓ^∞} to $\mathrm{Gal}(\overline{H}/H) \subseteq \mathrm{Gal}(\overline{K}/K)$, so it is clear that $\rho_{\ell^\infty}\left(\mathrm{Gal}(\overline{H}/H)\right)$ is a subgroup of $\rho_{\ell^\infty}\left(\mathrm{Gal}(\overline{K}/K)\right)$ of index dividing $h(K)$, the class number of K. Now as ρ_{ℓ^∞} factors through $\mathrm{Gal}(\overline{K}/K)$ we see that $\rho_{\ell^\infty}(K^\times)$ is trivial, so we can just consider the restriction of ρ_{ℓ^∞} to $\prod_{v\in\Omega_K} \mathcal{O}_{K,v}^\times$. We now remark that for an idèle $(a_v) \in \prod_{v\in\Omega_K} \mathcal{O}_{K,v}^\times$ theorem 5.5.1 implies

$$\varepsilon(a) = \prod_{v\in\Omega_K} \varepsilon\left(a_v\right) = \prod_{\substack{v:A \text{ has bad} \\ \text{reduction at v}}} \varepsilon(a_v) \in \mu(E),$$

whence $J := \ker\varepsilon \cap \prod_{v\in\Omega_K} \mathcal{O}_{K,v}^\times$ has index dividing $|\mu(E)|$ in $\prod_{v\in\Omega_K} \mathcal{O}_{K,v}^\times$, and likewise the index of $J_\ell := \ker\varepsilon \cap \prod_{v|\ell} \mathcal{O}_{K,v}^\times$ in $\prod_{v|\ell} \mathcal{O}_{K,v}^\times$ divides $|\mu(E)|$. Furthermore, since the function $a \mapsto \Phi_{(E,S)}\left(N_{K_\ell/E_\ell^*}(a_\ell)\right)^{-1}$ kills $\mathcal{O}_{K,v}^\times$ when $p_v \neq \ell$, we have $\rho_{\ell^\infty}(J) = \rho_{\ell^\infty}(J_\ell)$. Also notice that, upon restriction to J_ℓ, the representation ρ_{ℓ^∞} coincides with the map

$$\begin{aligned} \varphi_{\ell^\infty} : \prod_{v|\ell} \mathcal{O}_{K,v}^\times &\to (\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times \\ a &\mapsto \Phi_{(E,S)}\left(N_{K_\ell/E_\ell^*}(a)\right)^{-1}, \end{aligned}$$

and that if A has good reduction at v, then ρ_{ℓ^∞} and φ_{ℓ^∞} coincide on all of $\prod_{v|\ell} \mathcal{O}_{K,v}^\times$. For the sake of notational simplicity let us then set

$$\mu^* = \begin{cases} |\mu(E)|, \text{ if } A \text{ has bad reduction at} \\ \qquad\qquad \text{some place } v \text{ of characteristic } \ell \\ 1, \text{ otherwise} \end{cases}$$

We have proved:

Proposition 5.5.2. *For all primes ℓ the group G_{ℓ^∞} contains $\rho_{\ell^\infty}(J_\ell)$ as a subgroup of index dividing $|\mu(E)| \cdot h(K)$. We have $\rho_{\ell^\infty}(J_\ell) = \varphi_{\ell^\infty}(J_\ell)$, and if A has good reduction at all places v of characteristic ℓ we have $J_\ell = \prod_{v|\ell} \mathcal{O}_{K,v}^\times$. Finally,*

$$\left[\varphi_{\ell^\infty}\left(\prod_{v|\ell} \mathcal{O}_{K,v}^\times\right) : \rho_{\ell^\infty}(J_\ell)\right] \mid \mu^*. \tag{5.10}$$

We can now interpret $\varphi_{\ell\infty}$ as a map between algebraic tori: indeed, the norm N_{K/E^*} can be seen as a morphism $T_K \to T_{E^*}$, and $\prod_{v|\ell} \mathcal{O}_{K,v}^\times$ is nothing but $T_K(\mathbb{Z}_\ell)$, so the map $\varphi_{\ell\infty}$ is simply the map induced on $\mathbb{Z}_\ell$-points by

$$\left(\Phi_{(E,S)}\right)^{-1} \circ N_{K/E^*} : T_K \to \mathrm{MT}(A);$$

together with the previous proposition, this implies in particular that $\rho_{\ell\infty}(J_\ell) = \varphi_{\ell\infty}(J_\ell)$ is contained in $\mathrm{MT}(A)(\mathbb{Z}_\ell)$, and that $\varphi_{\ell\infty}(J_\ell)$ has index at most μ^* in $\varphi_{\ell\infty}(T_K(\mathbb{Z}_\ell))$. We thus want to understand the composition

$$T_K(\mathbb{Z}_\ell) \xrightarrow{N_{K/E^*}} T_{E^*}(\mathbb{Z}_\ell) \xrightarrow{\psi_\ell} \mathrm{MT}(A)(\mathbb{Z}_\ell),$$

where for simplicity of notation we write ψ_ℓ for the base-change to $\mathbb{Q}_\ell$ of the map $\left(\Phi_{(E,S)}(\cdot)\right)^{-1}$. Even though the extension K/E^* is in general non-abelian, the cokernel of N_{K/E^*} can be understood through class field theory:

Theorem 5.5.3. *([4, Theorem 7 on p. 161]) Let L/M be an extension of local fields, and let L_{ab} be the largest abelian subextension of L/M. Then we have $N_{L/M}L^\times = N_{L_{ab}/M}\left(L_{ab}^\times\right)$, and the cokernel $\dfrac{M^\times}{N_{L/M}L^\times}$ has order dividing $[L:M]$.*

Notice that the image of ψ_ℓ is open and $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ is compact, hence the cokernel of

$$\psi_\ell : T_{E^*}(\mathbb{Z}_\ell) \xrightarrow{\psi_\ell} \mathrm{MT}(A)(\mathbb{Z}_\ell)$$

is finite; since furthermore by theorem 5.5.3 $\left|\dfrac{T_{E^*}(\mathbb{Z}_\ell)}{N_{K/E^*}(T_K(\mathbb{Z}_\ell))}\right|$ divides $[K:E^*]$ we find that

$$[\mathrm{MT}(A)(\mathbb{Z}_\ell) : \varphi_{\ell\infty}(T_K(\mathbb{Z}_\ell))] \text{ divides } [K:E^*] \cdot \left|\frac{\mathrm{MT}(A)(\mathbb{Z}_\ell)}{\psi_\ell\left(T_{E^*}(\mathbb{Z}_\ell)\right)}\right|. \tag{5.11}$$

Remark 5.5.4. When ℓ is unramified in K the local norm $T_K(\mathbb{Z}_\ell) \to T_{E^*}(\mathbb{Z}_\ell)$ is surjective and the factor $[K:E^*]$ can be omitted, cf. [115, Corollary to Proposition 3 of Chapter V].

It is clear that $\psi_\ell = \Phi_{(E,S)}^{-1}$ and $\Phi_{(E,S)}$ have the same cokernel, so ultimately we just need to compute the cokernel of the reflex norm. Denote by T' the kernel of $\Phi_{(E,S)}$ and write F for the torsion of its character group $\widehat{T'}$. By proposition 5.2.15 we have $|F| \leq \lfloor 2^{-r}(r+1)^{(r+1)/2} \rfloor$, where $r = \dim \mathrm{Im}\, \Phi_{(E,S)}^* = \mathrm{rk}\, \mathrm{MT}(A)$ does not exceed $g+1$. Set now $T = T_{E^*} \otimes \mathbb{Q}_\ell$ and $T'' = \mathrm{MT}(A) \otimes \mathbb{Q}_\ell$, and let L be one of the fields appearing in the decomposition of $\tilde{E} \otimes \mathbb{Q}_\ell$ as a direct sum of fields: $L/\mathbb{Q}_\ell$ is then a finite Galois extension that splits T (recall that $\tilde{E}$ is Galois and contains E^*). If ℓ is unramified in E (hence in $\tilde{E}$) the extension $L/\mathbb{Q}_\ell$ is itself unramified, so T has good reduction over $\mathbb{Q}_\ell$; furthermore, $[L:\mathbb{Q}_\ell] \mid [\tilde{E}:\mathbb{Q}] \mid 2^g \cdot g!$ (cf. lemma 5.2.6).

Applying proposition 5.3.1 to the surjection of algebraic tori $T \xrightarrow{\Phi_{(E,S)}} T''$ we find that

$$\left|\mathrm{coker}\left(T_{E^*}(\mathbb{Z}_\ell) \xrightarrow{\Phi_{(E,S)}} \mathrm{MT}(A)(\mathbb{Z}_\ell)\right)\right| \text{ divides } |F| \cdot |F|_\ell^{-[L:\mathbb{Q}_\ell]}, \tag{5.12}$$

and the right hand side in turn divides $|F| \cdot |F|_\ell^{-2^g g!}$; we have thus almost completely established the following result:

Theorem 5.5.5. *(Theorem 5.1.3) Let A/K be an absolutely simple abelian variety of dimension g admitting complex multiplication over K by the CM type (E,S), and let ℓ be a prime number. If A*

has bad reduction at a place of K dividing ℓ let $\mu^* = |\mu(E)|$, the number of roots of unity in E; if on the contrary A has good reduction at all places of K of characteristic ℓ set $\mu^* = 1$. Denote by r the rank of $\mathrm{MT}(A)$ and by F the group of connected components of the kernel of the reflex norm $T_{E^*} \to T_E$, where E^* is the reflex field of E. Then:

(1) The index $[G_{\ell^\infty} : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ does not exceed $|\mu(E)| \cdot h(K)$, where $h(K)$ is the class number of K.

(2) We have $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)] \leq \mu^* \cdot [K : E^*] \cdot |F|^{2r}$.

(3) If ℓ is unramified in E and does not divide $|F|$, then the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ divides $\mu^* \cdot [K : E^*] \cdot |F|$. If ℓ is also unramified in K, the bound can be improved to $\mu^* \cdot |F|$.

Finally we have $r \leq g + 1$ and $|F| \leq f(r) \leq f(g + 1)$, where

$$f(x) = \left\lfloor 2 \left(\frac{x+1}{4} \right)^{(x+1)/2} \right\rfloor.$$

Proof. We have already proved (1): the intersection $G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)$ contains $\varphi_{\ell^\infty}(J_\ell) = \rho_{\ell^\infty}(J_\ell)$, and by proposition 5.5.2 the group $\varphi_{\ell^\infty}(J_\ell)$ has index at most $|\mu(E)| \cdot h(K)$ in G_{ℓ^∞}. As for part (2), the exact sequence

$$1 \to T' \to T_{E^*} \otimes \mathbb{Q}_\ell \to \mathrm{MT}(A) \otimes \mathbb{Q}_\ell \to 1$$

induces, by quotienting out by $(T')^0$ (the connected component of the identity of T'), the exact sequence

$$1 \to \mathcal{F} \to \frac{T_{E^*} \otimes \mathbb{Q}_\ell}{(T')^0} \xrightarrow{\tau_\ell} \mathrm{MT}(A) \otimes \mathbb{Q}_\ell \to 1,$$

where $\mathcal{F}$ is a finite group scheme of order $|F|$. Proposition 5.4.1 implies

$$\begin{aligned} \left| \frac{\mathrm{MT}(A)(\mathbb{Z}_\ell)}{\psi_\ell\left(T_{E^*}(\mathbb{Z}_\ell)\right)} \right| &= \left| \mathrm{coker}\left(\tau_\ell : \frac{T_{E^*} \otimes \mathbb{Q}_\ell}{(T')^0}(\mathbb{Z}_\ell) \to \mathrm{MT}(A)(\mathbb{Z}_\ell) \right) \right| \\ &\leq |\deg(\tau_\ell)|^{\dim \mathrm{MT}(A)} |\deg(\tau_\ell)|_\ell^{-\dim \mathrm{MT}(A)} \\ &= |F|^{\dim \mathrm{MT}(A)} |F|_\ell^{-\dim \mathrm{MT}(A)}, \end{aligned}$$

which, together with equations (5.10) and (5.11), gives the desired result. Finally, consider part (3). As $\rho_{\ell^\infty}(J_\ell)$ is a subgroup of $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ divides $[MT(A)(\mathbb{Z}_\ell) : \rho_{\ell^\infty}(J_\ell)]$, and we can write

$$\begin{aligned} \left| \frac{\mathrm{MT}(A)(\mathbb{Z}_\ell)}{\rho_{\ell^\infty}(J_\ell)} \right| &\mid \mu^* \cdot [\mathrm{MT}(A)(\mathbb{Z}_\ell) : \varphi_{\ell^\infty}(T_K(\mathbb{Z}_\ell))] && \text{(by (5.10))} \\ &\mid \mu^* \cdot [K : E^*] \cdot |\mathrm{coker}\,(\psi_\ell : T_{E^*}(\mathbb{Z}_\ell) \to \mathrm{MT}(A)(\mathbb{Z}_\ell))| && \text{(by (5.11))} \\ &\mid \mu^* \cdot [K : E^*] \cdot |F| \cdot |F|_\ell^{-2^g g!}. && \text{(by (5.12))} \end{aligned}$$

Since by assumption ℓ does not divide $|F|$ we conclude that $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)]$ divides $\mu^* \cdot [K : E^*] \cdot |F|$. Finally, when ℓ is unramified in K the factor $[K : E^*]$ can be omitted, cf. remark 5.5.4. □

Starting from equations (5.11) and (5.12) it is also easy to prove the following result, which might have some independent interest:

Proposition 5.5.6. *(Proposition 5.1.6) Let A/K be an absolutely simple abelian variety admitting complex multiplication (over K) by the CM field E, and let ℓ be a prime unramified in E. Let E^* be the reflex field of E and suppose that A has good reduction at all places of K of characteristic ℓ.*

- *The index $[\mathrm{MT}(A)(\mathbb{F}_\ell) : G_\ell \cap \mathrm{MT}(A)(\mathbb{F}_\ell)]$ divides $[K : E^*] \cdot |F|$.*
- *If ℓ is also unramified in K, then $[\mathrm{MT}(A)(\mathbb{F}_\ell) : G_\ell \cap \mathrm{MT}(A)(\mathbb{F}_\ell)]$ divides $|F|$.*

Proof. By proposition 5.2.2 the hypothesis implies that T_{E^*} has good reduction at ℓ, hence the same is true for its quotient $\mathrm{MT}(A)$, which therefore defines a torus over $\mathbb{F}_\ell$: in particular, the group $\mathrm{MT}(A)(\mathbb{F}_\ell)$ makes sense and its order is not divisible by ℓ. On the other hand, the index of $G_\ell \cap \mathrm{MT}(A)(\mathbb{F}_\ell)$ in $\mathrm{MT}(A)(\mathbb{F}_\ell)$ divides $[K : E^*] \cdot |F| \cdot |F|_\ell^{-2^g g!}$ by proposition 5.5.2 and equations (5.11) and (5.12), and since $|\mathrm{MT}(A)(\mathbb{F}_\ell)|$ is prime to ℓ we deduce that $[\mathrm{MT}(A)(\mathbb{F}_\ell) : G_\ell \cap \mathrm{MT}(A)(\mathbb{F}_\ell)]$ divides $[K : E^*] \cdot |F|$ as claimed. The second part follows by the same argument and remark 5.5.4. □

5.6 The Mumford-Tate group in the nondegenerate case

In this section we consider CM abelian varieties A with nondegenerate CM type, that is to say we assume that $\mathrm{rank}(\mathrm{MT}(A)) = \dim A + 1$: this is the "generic" case, and it is also known that all simple CM varieties of prime dimension have nondegenerate CM type (a result due to Ribet, cf. [111]). In this situation we have the following bounds on the order of $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$:

Theorem 5.6.1. *Suppose A is simple of nondegenerate CM type. For all primes $\ell \neq 2$ and all $n \geq 1$ we have*

$$(1 - 1/\ell)^{g+1} \cdot \ell^{(g+1)n} \leq |\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \leq 2^g (1 + 1/\ell)^{g-1} \ell^{(g+1)n},$$

while for $\ell = 2$ and all $n \geq 1$ we have

$$\frac{1}{2^{2g+3}} \cdot 2^{(g+1)n} \leq |\mathrm{MT}(A)(\mathbb{Z}/2^n\mathbb{Z})| \leq \frac{1}{2} 4^g \cdot 2^{(g+1)n}.$$

The proof of this result will occupy sections 5.6.1 and 5.6.2, while in sections 5.6.3 and 5.6.4 we discuss the special cases of elliptic curves and abelian surfaces.

5.6.1 The natural filtration on the norm-1 torus

Let $\ell \neq 2$ be a rational prime, L be a finite extension of $\mathbb{Q}_\ell$ and τ be an involution of L. Denote L^τ the fixed field of τ, so that L/L^τ is a quadratic (Galois) extension. Fix a squarefree $d \in \mathcal{O}_{L^\tau}$ such that $L = L^\tau\left(\sqrt{d}\right)$ and consider the (multiplicative) group

$$C = \left\{x \in \mathcal{O}_L^\times \;\middle|\; x \cdot \tau(x) = 1\right\}.$$

We write λ for a uniformizer of L^τ, set $e = e\left(L^\tau/\mathbb{Q}_\ell\right)$, and consider v_ℓ and v_λ as valuations on $\overline{\mathbb{Q}_\ell}$, normalized so as to have $v_\lambda(\lambda) = 1$ and $v_\ell(\ell) = 1$; in particular, $v_\lambda = e \cdot v_\ell$. We want to investigate the structure of the filtration of C given by $C(n) := \left\{x \in C \;\middle|\; v_\lambda(x - 1) \geq n\right\}$. It is easy to see that every $x \in C(1)$ can be represented as

$$x = 1 + 2du \cdot \lambda^{2+2v} + 2u_2 \cdot \lambda^{1+v}\sqrt{d}$$

with $u, u_2 \in \mathcal{O}_{L^\tau}^\times$ and $v \in \mathbb{N}$ subject to the condition

$$u(1 + du \cdot \lambda^{2+2v}) = u_2^2. \tag{5.13}$$

Furthermore, for $n \geq 1$ we have an exact sequence of abelian groups

$$\begin{array}{ccccccc} 0 & \to & C(n+1) & \to & C(n) & \xrightarrow{\alpha} & \mathcal{O}_L/(\lambda)\mathcal{O}_L, \\ & & & & x & \mapsto & \left[\dfrac{x-1}{2\lambda^n}\right] \end{array}$$

where $[\cdot]$ denotes the class of an element of $\mathcal{O}_L$ in the quotient $\mathcal{O}_L/(\lambda)\mathcal{O}_L$. Let us describe the image of α for $n \geq 1$. Clearly $x \in C(n)$ implies $v \geq n-1$, and for $v \geq n$ we have $\alpha(x) = 0$; when $v = n-1$ we have $\alpha(x) = [u_2\sqrt{d}]$. Notice now that we have an injection (of additive groups) $\frac{\mathcal{O}_{L^\tau}}{(\lambda)\mathcal{O}_{L^\tau}} \hookrightarrow \frac{\mathcal{O}_L}{(\lambda)\mathcal{O}_L}$ induced by $x \mapsto x\sqrt{d}$, and we claim that all points in the image of this embedding can be realized as $\alpha(x)$ for some $x \in C(n)$. This is clear for the zero element, so let us consider an element of the form $[u_2\sqrt{d}]$ with $u_2 \in \mathcal{O}_{L^\tau}^\times$. Consider the equation

$$t\left(1 + \lambda^{2n} dt\right) = u_2^2 \tag{5.14}$$

in the variable t. By Hensel's lemma, the discriminant $\Delta := 1 + 4u_2^2\lambda^{2n}d$ is a square in $\mathcal{O}_{L^\tau}$ (recall that $n > 0$). Let $1 + z$ be the square root of Δ that is congruent to 1 modulo λ: then z satisfies $(1+z)^2 = 1 + 4u_2^2\lambda^{2n}d$, from which we easily find $v_\lambda(z) = 2n + v_\lambda(d)$. It follows that $u := \dfrac{-1+\sqrt{\Delta}}{2d\lambda^{2n}} = \dfrac{z}{2d\lambda^{2n}}$ is a solution to equation (5.14) which is also a λ-adic unit. We can then set $x = 1 + 2du \cdot \lambda^{2n} + 2u_2 \cdot \lambda^n\sqrt{d}$: by construction x is an element of $C(n)$, and it satisfies $\alpha(x) = [u_2\sqrt{d}]$. This shows that the image of α is in bijection with $\frac{\mathcal{O}_{L^\tau}}{(\lambda)\mathcal{O}_{L^\tau}}$. Finally, essentially the same argument can be repeated when $\ell = 2$, except that Hensel's lemma is now only applicable for $n > v_\lambda(2)$. We thus deduce the following lemma:

Lemma 5.6.2. *Suppose $\ell \neq 2$. For every $n \geq 1$, the quotient $C(n)/C(n+1)$ has order $\left|\frac{\mathcal{O}_{L^\tau}}{(\lambda)\mathcal{O}_{L^\tau}}\right|$. For $\ell = 2$ the same conclusion holds for every $n > v_\lambda(2)$.*

The quotients $C(n)/C(n+1)$ for small values of n are described by the following lemma:

Lemma 5.6.3. *Let f be the inertia degree of L^τ over $\mathbb{Q}_\ell$. Suppose first $\ell \neq 2$: then the quotient $\frac{C(0)}{C(1)}$ has order either $2\ell^f$ or $\ell^f + 1$, with the first (resp. second) case happening exactly when L/L^τ is ramified (resp. unramified). Suppose on the other hand that $\ell = 2$ and $n \leq v_\lambda(2)$: then the quotient $\frac{C(n)}{C(n+1)}$ has order at most 4^f.*

Before giving a proof, recall the following

Definition 5.6.4. Let L be a finite extension of $\mathbb{Q}_\ell$ with ring of integers $\mathcal{O}_L$ and residue field $\mathbb{F}$. Let $\pi : \mathcal{O}_L \to \mathbb{F}$ be the canonical projection. The **Teichmüller lift** is the unique group homomorphism $\omega : \mathbb{F}^\times \to \mathcal{O}_L^\times$ such that, for all $y \in \mathbb{F}^\times$, the element $\omega(y) \in \mathcal{O}_L^\times$ is the unique solution to the equation $x^{|\mathbb{F}|-1} = 1$ satisfying $\pi(x) = y$.

Proof. Consider first the case of L/L^τ being unramified (and $\ell \neq 2$). Let $\pi : \mathcal{O}_L \to \mathbb{F} := \frac{\mathcal{O}_L}{(\lambda)\mathcal{O}_L}$ be the canonical projection, and observe that $\mathbb{F}$ has order ℓ^{2f}. It is clear that π restricts to a map $C(0) \to \mathbb{F}^\times$, and on the other hand $x \in C(0)$ maps to 1 if and only if $v_\lambda(x-1) > 0$, i.e. if and only if $x \in C(1)$: this implies that $C(0)/C(1)$ injects into $\mathbb{F}^\times$. The involution τ induces on $\mathbb{F}$ an

automorphism $\tau_{\mathbb{F}}$, which is necessarily the unique nontrivial involution $x \mapsto x^{\ell^f}$. Let now $x \in C(0)$. By definition we have $x \cdot \tau(x) = 1$, hence $1 = \pi(x) \cdot \pi(\tau(x)) = \pi(x) \cdot \tau_{\mathbb{F}}(\pi(x)) = \pi(x)^{\ell^f+1}$, so $C(0)/C(1)$ injects into the subgroup H of $\mathbb{F}^\times$ given by the roots of unity of order dividing $\ell^f + 1$. The group H is of order $\ell^f + 1$, and it is not hard to see that $C(0)/C(1)$ surjects onto it: indeed for every $h \in H$ we have $\omega(h) \in C(0)$, and by definition $\pi(\omega(h)) = h$. Suppose on the other hand that L/L^τ is ramified, so that $L = L^\tau(\sqrt{d})$ with $v_\lambda(d) = 1$. Again we see that $C(0)/C(1)$ injects into $\mathbb{F}^\times := \left(\frac{\mathcal{O}_L}{(\lambda)\mathcal{O}_L}\right)^\times$ (which however is not a field anymore), and the involution τ acts on an element $[a + b\sqrt{d}] \in \left(\frac{\mathcal{O}_L}{(\lambda)\mathcal{O}_L}\right)^\times$, with $a, b \in \mathcal{O}_{L^\tau}$, by sending it to $[a - b\sqrt{d}]$. Writing $\pi(x) = [a + b\sqrt{d}]$, the equation $x\tau(x) = 1$ implies $[a^2 - db^2] = 1$, which in turn, since $v_\lambda(d) = 1$, means $[a^2] = 1$ and $[a] = \pm 1$. This shows that $C(0)/C(1)$ injects into $\{\pm 1\} \times \frac{\mathcal{O}_{L^\tau}}{(\lambda)\mathcal{O}_{L^\tau}}$, a set with $2 \cdot \ell^f$ elements. On the other hand, for any value of $[\pm 1 + b\sqrt{d}] \in \mathbb{F}^\times$, the equation $a^2 = 1 + db^2$ (with fixed b, in the variable a) admits solutions in $\mathcal{O}_{L^\tau}$ by Hensel's lemma; the elements $\pm a + b\sqrt{d} \in C(0)$ then satisfy

$$(\pm a + b\sqrt{d}) \cdot \tau(\pm a + b\sqrt{d}) = a^2 - db^2 = 1,$$

and on the other hand $\pi(\pm a + b\sqrt{d}) = [\pm 1 + b\sqrt{d}]$, so $C(0)/C(1)$ actually projects surjectively on $\{\pm 1\} \times \frac{\mathcal{O}_{L^\tau}}{(\lambda)\mathcal{O}_{L^\tau}}$; this shows that $|C(0)/C(1)| = 2\ell^f$ as claimed. The upper bound for $\ell = 2$ likewise follows from the fact that for any $n \geq 0$ the quotient $C(n)/C(n+1)$ injects into $\frac{\mathcal{O}_L}{(\lambda)\mathcal{O}_L}$. □

5.6.2 The order of $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$

Let E be a CM field of degree $2g$ over $\mathbb{Q}$ and T_E be the associated algebraic torus, and let τ denote complex conjugation on E. If A/K is an abelian variety with complex multiplication by the nondegenerate CM type (E, S), it is known that we have

$$\mathrm{MT}(A)(B) = \left\{x \in (E \otimes_{\mathbb{Q}} B)^\times \;\middle|\; x\tau(x) \in B^\times\right\} \quad \forall\, \mathbb{Q}\text{-algebra } B.$$

We can also consider the 'norm 1' (or Hodge) subtorus of $\mathrm{MT}(A)$ given as a functor by

$$\mathrm{Hg}(A)(B) = \left\{x \in (E \otimes_{\mathbb{Q}} B)^\times \;\middle|\; x\tau(x) = 1\right\} \quad \forall\, \mathbb{Q}\text{-algebra } B.$$

We aim to give bounds on the number of $\frac{\mathbb{Z}}{\ell^n\mathbb{Z}}$-points of $\mathrm{MT}(A)$, but it is easier to first consider $\mathrm{Hg}(A)$. If we write $E \otimes \mathbb{Q}_\ell \cong \prod_{i=1}^s F_i$ (a product of fields), we have

$$\mathrm{Hg}(A)(\mathbb{Q}_\ell) = \left\{x = (x_1, \ldots, x_s) \in \prod_{i=1}^s F_i^\times \;\middle|\; x\tau(x) = 1\right\}.$$

We can renumber the F_i's in such a way that τ acts by exchanging F_{2i-1} and F_{2i} for $i = 1, \ldots, r$ and it acts as an involution on F_i for $i = 2r+1, \ldots, s$.
With this convention, a point $(x_1, \ldots, x_{2r}, x_{2r+1}, \ldots, x_s) \in \prod_{i=1}^s F_i^\times$ is in $\mathrm{Hg}(A)(\mathbb{Q}_\ell)$ if and only if $x_{2i-1}x_{2i} = 1$ for $i = 1, \ldots, r$ and $x_i\tau(x_i) = 1$ for $i = 2r+1, \ldots, s$, that is,

$$\mathrm{Hg}(A)(\mathbb{Q}_\ell) \cong \prod_{i=1}^r \left\{x_{2i-1} \in F_{2i-1}^\times\right\} \times \prod_{i=2r+1}^s \left\{x_i \in F_i^\times \;\middle|\; x_i\tau(x_i) = 1\right\}. \tag{5.15}$$

The character groups of $\mathrm{MT}(A)_{\mathbb{Q}_\ell}$ and of $\mathrm{Hg}(A)_{\mathbb{Q}_\ell}$ are quotients of $\widehat{T_{E,\mathbb{Q}_\ell}}$, which in turn is generated by the elements of the form $(\chi_1, \ldots, \chi_s)$, where χ_i ranges over the embeddings of F_i in $\overline{\mathbb{Q}_\ell}$. It follows that a point $x \in \mathrm{Hg}(A)(\mathbb{Q}_\ell)$ is in $\mathrm{Hg}(A)(\mathbb{Z}_\ell)$ if and only if for any choice of embeddings $\chi_i : F_i \hookrightarrow \overline{\mathbb{Q}_\ell}$

we have $\chi_i(x_i) \in \mathcal{O}_{\chi_i(F_i)}$; as the property of being ℓ-integral is Galois-invariant we deduce that a necessary and sufficient condition is $x_i \in \mathcal{O}_{F_i}^\times$. Hence we find

$$\mathrm{Hg}(A)(\mathbb{Z}_\ell) \cong \prod_{i=1}^{r} \mathcal{O}_{F_{2i-1}}^\times \times \prod_{i=2r+1}^{s} \left\{ x_i \in \mathcal{O}_{F_i}^\times \mid x_i \tau(x_i) = 1 \right\},$$

and a perfectly analogous argument shows that

$$\mathrm{Hg}(A)(1+\ell^n\mathbb{Z}_\ell) \cong \prod_{i=1}^{r} \left\{ x_{2i-1} \in \mathcal{O}_{F_{2i-1}}^\times \mid v_\ell(x_{2i-1}-1) \geq n \right\} \times \\ \times \prod_{i=2r+1}^{s} \left\{ x_i \in \mathcal{O}_{F_i}^\times \mid v_\ell(x_i - 1) \geq n,\ x_i\tau(x_i) = 1 \right\}.$$

Write e_i and f_i for the ramification index and inertia degree of F_i^τ over $\mathbb{Q}_\ell$, and λ_i for a uniformizer of F_i^τ $(i = 2r+1, \dots, s)$. The order of $\left| \frac{\mathrm{Hg}(A)(\mathbb{Z}_\ell)}{\mathrm{Hg}(A)(1+\ell^n\mathbb{Z}_\ell)} \right|$ is then given by

$$|\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| = \prod_{i=1}^{r} \left| \frac{\mathcal{O}_{F_{2i-1}}^\times}{1+\ell^n\mathcal{O}_{F_{2i-1}}} \right| \times \prod_{i=2r+1}^{s} \left| \frac{C^{(i)}(0)}{C^{(i)}(ne_i)} \right|, \tag{5.16}$$

where

$$C^{(i)}(k) = \left\{ x_i \in \mathcal{O}_{F_i}^\times \mid v_{\lambda_i}(x_i - 1) \geq k,\ x_i \cdot \tau(x_i) = 1 \right\}$$

is the filtration we studied in the previous section for the field F_i and the involution $\tau|_{F_i}$. For $i = 1, \dots, r$ let furthermore π_i (resp. e_i, f_i) be a uniformizer (resp. the ramification index over $\mathbb{Q}_\ell$, the inertia degree over $\mathbb{Q}_\ell$) of F_{2i-1}. We now compute the order of $\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$. Basic properties of local fields show that $\left| \frac{\mathcal{O}_{F_{2i-1}}^\times}{1+\ell^n\mathcal{O}_{F_{2i-1}}} \right|$ has order

$$\left| \frac{\mathcal{O}_{F_{2i-1}}^\times}{1+\pi_i\mathcal{O}_{F_{2i-1}}} \right| \cdot \prod_{j=1}^{ne_i-1} \left| \frac{1+(\pi_i)^j\mathcal{O}_{F_{2i-1}}}{1+(\pi_i)^{j+1}\mathcal{O}_{F_{2i-1}}} \right| = \left(\ell^{f_i} - 1\right) \cdot \ell^{f_i(ne_i-1)},$$

while (for $\ell \neq 2$) lemma 5.6.2 gives

$$\left| \frac{C^{(i)}(0)}{C^{(i)}(ne_i)} \right| = \left| \frac{C^{(i)}(0)}{C^{(i)}(1)} \right| \cdot \left| \frac{C^{(i)}(1)}{C^{(i)}(ne_i)} \right| = \left| \frac{C^{(i)}(0)}{C^{(i)}(1)} \right| \cdot \ell^{f_i(ne_i-1)}.$$

Now notice that $s - 2r$ does not exceed g: indeed $[F_i : F_i^\tau] = 2$ for every $i = 2r+1, \dots, r$, hence $2g = [E \otimes \mathbb{Q}_\ell : \mathbb{Q}_\ell] \geq \sum_{i=2r+1}^{s} [F_i : \mathbb{Q}_\ell] \geq 2(s-2r)$ as claimed. Applying lemma 5.6.3 we then deduce that the order of $\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ is at most

$$\begin{aligned} &\prod_{i=1}^{r} \ell^{nf_ie_i} \cdot \prod_{i=2r+1}^{s} 2\left(1+\ell^{f_i}\right)\left(\ell^{f_i}\right)^{ne_i-1} \\ &\quad = 2^{s-2r} \prod_{i=1}^{2r} \ell^{\frac{1}{2}n[F_i:\mathbb{Q}_\ell]} \prod_{i=2r+1}^{s} \left(1+\ell^{-f_i}\right)\left(\ell^{f_i}\right)^{ne_i} \\ &\quad \leq 2^{s-2r}\,(1+1/\ell)^{s-2r} \prod_{i=1}^{s} \ell^{\frac{1}{2}n[F_i:\mathbb{Q}_\ell]} \\ &\quad \leq 2^g\,(1+1/\ell)^g\,\ell^{gn}, \end{aligned}$$

and at least

$$\begin{aligned}\prod_{i=1}^{r}\left(\ell^{f_i}-1\right)\ell^{(ne_i-1)f_i}\prod_{i=2r+1}^{s}(\ell^{f_i}+1)\left(\ell^{f_i}\right)^{ne_i-1}&\\ \geq (1-1/\ell)^r\cdot\prod_{i=1}^{r}\ell^{nf_ie_i}\prod_{i=2r+1}^{s}\ell^{nf_ie_i}&\\ =(1-1/\ell)^r\cdot\prod_{i=1}^{2r}\ell^{\frac{1}{2}n[F_i:\mathbb{Q}_\ell]}\times\prod_{i=2r+1}^{s}\ell^{\frac{1}{2}n[F_i:\mathbb{Q}_\ell]}&\\ \geq (1-1/\ell)^g\cdot\ell^{gn};&\end{aligned}$$

moreover, if for at least one index $i\in\{2r+1,\dots,s\}$ the extension F_i/F_i^τ is ramified, then we see from lemma 5.6.3 that the lower bound can be improved to

$$|\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})|\geq 2(1-1/\ell)^g\cdot\ell^{gn}. \tag{5.17}$$

To finish the proof of theorem 5.6.1 we shall use the following lemma:

Lemma 5.6.5. *Consider the map*

$$\begin{array}{rccc}\Psi: & \mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})\times(\mathbb{Z}/\ell^n\mathbb{Z})^\times & \to & \mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})\\ & (h,m) & \mapsto & m^{-1}h.\end{array}$$

If $\ell\neq 2$, the group $\operatorname{Im}\Psi$ has order equal to $\frac{1}{2}|\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})|\times(1-1/\ell)\ell^n$ and has index at most 2 in $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$. Moreover, Ψ is surjective if and only if for all $x\in\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ the number $x\tau(x)$ is a square in $(\mathbb{Z}/\ell^n\mathbb{Z})^\times$. On the other hand, for $\ell=2$ we have

- *for $n=1$, the group $\operatorname{Im}\Psi$ has order equal to that of $|\mathrm{Hg}(A)(\mathbb{Z}/2\mathbb{Z})|$ and Ψ is surjective;*
- *for $n=2$, the group $\operatorname{Im}\Psi$ has order equal to that of $|\mathrm{Hg}(A)(\mathbb{Z}/4\mathbb{Z})|$ and $\operatorname{Im}\Psi$ has index either 1 or 2 in $\mathrm{MT}(A)(\mathbb{Z}/4\mathbb{Z})$;*
- *for $n\geq 3$, the group $\operatorname{Im}\Psi$ has order equal to $2^{n-3}\cdot|\mathrm{Hg}(A)(\mathbb{Z}/2^n\mathbb{Z})|$ and $\operatorname{Im}\Psi$ has index 1, 2 or 4 in $\mathrm{MT}(A)(\mathbb{Z}/2^n\mathbb{Z})$;*

Proof. Let us start with the case $\ell\neq 2$. The kernel of ψ is given by the intersection of $\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ and $(\mathbb{Z}/\ell^n\mathbb{Z})^\times$ inside $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$, namely

$$\left\{h\in(\mathbb{Z}/\ell^n\mathbb{Z})^\times\;\middle|\;h\tau(h)=1\right\}=\left\{h\in(\mathbb{Z}/\ell^n\mathbb{Z})^\times\;\middle|\;h^2=1\right\}=\{\pm 1\},$$

so $\operatorname{Im}\Psi$ has order

$$\frac{1}{|\ker\Psi|}\cdot|\,\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})|\cdot\left|(\mathbb{Z}/\ell^n\mathbb{Z})^\times\right|=\frac{(\ell-1)\ell^{n-1}}{2}\cdot|\,\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})|$$

as claimed.

As for the index of $\operatorname{Im}\Psi$, notice first that for every $x=m^{-1}h\in\operatorname{Im}\Psi$ we have that $x\cdot\tau(x)=m^{-2}$ is a square in $(\mathbb{Z}/\ell^n\mathbb{Z})$, so if Ψ is surjective we necessarily have $x\cdot\tau(x)\in(\mathbb{Z}/\ell^n\mathbb{Z})^{\times 2}$ for every $x\in\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$. Conversely, suppose that for every x in $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ the number $x\tau(x)$ is a square in $(\mathbb{Z}/\ell^n\mathbb{Z})^\times$, say $x\tau(x)=\mu(x)^2$ with $\mu(x)\in(\mathbb{Z}/\ell^n\mathbb{Z})^\times$. Then every x can be written as $x=\mu(x)\cdot\frac{x}{\mu(x)}$, and since $\frac{x}{\mu(x)}$ is in $\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ this shows that x belongs to $\operatorname{Im}\Psi$, which is therefore surjective.

Finally, if there is a $y \in \mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ such that $y\tau(y)$ is not a square in $(\mathbb{Z}/\ell^n\mathbb{Z})^\times$, then using the fact that $(\mathbb{Z}/\ell^n\mathbb{Z})^{\times 2}$ is of index 2 in $(\mathbb{Z}/\ell^n\mathbb{Z})^\times$ we easily see that for every $x \in \mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ either x or xy belongs to $\mathrm{Im}\,\Psi$, thus proving the remaining claim. The conclusion for $\ell = 2$ follows by the same argument upon noticing that $\frac{(\mathbb{Z}/2^n\mathbb{Z})^\times}{(\mathbb{Z}/2^n\mathbb{Z})^{\times 2}}$ has order $1, 2$, or 4, according to whether n is 1, 2, or at least 3. □

Combining this last lemma with our previous estimates gives the desired upper bound

$$\begin{aligned}|\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \leq 2\,|\mathrm{Im}\,\Psi| &= |\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \times \left|(\mathbb{Z}/\ell^n\mathbb{Z})^\times\right| \\ &\leq 2^g\,(1+1/\ell)^{g-1}\,\ell^{(g+1)n}.\end{aligned}$$

As for the lower bound, suppose first that for at least one index i in the set $\{2r+1, \ldots, s\}$ the extension L_i/L_i^τ is ramified: then using the lower bound of equation (5.17) (which is conditional on this hypothesis) we find

$$|\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \geq \frac{1}{2}\,|\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \times \left|(\mathbb{Z}/\ell^n\mathbb{Z})^\times\right| \geq (1-1/\ell)^{g+1}\ell^{(g+1)n}.$$

Suppose on the other hand that L_i/L_i^τ is unramified for every $i = 2r+1, \ldots, s$: then we claim that map Ψ from lemma 5.6.5 is not surjective. Assuming this is the case, we have

$$|\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \geq 2 \times \frac{1}{2} \times |\mathrm{Hg}(A)(\mathbb{Z}/\ell^n\mathbb{Z})| \times \left|(\mathbb{Z}/\ell^n\mathbb{Z})^\times\right| \geq (1-1/\ell)^{g+1}\ell^{(g+1)n},$$

which is what we want to show. We are thus reduced to proving that Ψ is not surjective, or equivalently (by lemma 5.6.5), to showing that there is an $x \in \mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ such that $x\tau(x)$ is not a square in $(\mathbb{Z}/\ell^n\mathbb{Z})^\times$. By the same argument that leads to equations (5.15) and (5.16), we can represent elements of $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ as tuples

$$(x_1, \ldots, x_{2r}, x_{2r+1}, \ldots, x_s, m) \in \prod_{i=1}^{2r} \mathcal{O}_{F_i}^\times \times \prod_{j=2r+1}^{s} \mathcal{O}_{F_j}^\times \times \mathbb{Z}_\ell^\times,$$

satisfying $x_{2i-1}x_{2i} = m$ for $i = 1, \ldots, r$ and $x_j\tau(x_j) = m$ for $j = 2r+1, \ldots, s$. Now if $2r = s$ it is clear that $\mathrm{MT}(A)(\mathbb{Z}_\ell/\ell^n\mathbb{Z})$ contains elements x such that $x\tau(x)$ is not a square (it suffices to choose $m \in \mathbb{Z}_\ell^\times$ which is not a square in $\mathbb{Z}/\ell^n\mathbb{Z}$ and set $x_{2i-1} = 1, x_{2i} = m$ for $i = 1, \ldots, r$), so we can assume $s > 2r$. For $j = 2r+1, \ldots, s$ write $F_j = F_j^\tau(\sqrt{d_j})$ for some squarefree $d_j \in \mathcal{O}_{F_j}^\times$ (recall that we assume F_j/F_j^τ to be unramified), and likewise write $x_j = a_j + b_j\sqrt{d_j}$ for some $a_j, b_j \in \mathcal{O}_{F_j^\tau}$. We claim that since F_j/F_j^τ is unramified every element $m \in \mathbb{Z}_\ell^\times$ can be represented as $a_j^2 - d_j b_j^2$ for some choice of $a_j, b_j \in \mathcal{O}_{F_j^\tau}$. To see this, notice that for fixed m and d_j the conic section $\mathcal{C} : \{a^2 - d_j b^2 = mc^2\}$ admits a point (a_0, b_0, c_0) over the residue field of F_j^τ; as d_j is not a square in F_j^τ we cannot have $c_0 = 0$, and since $\mathcal{C}$ is smooth the point (a_0, b_0, c_0) lifts to a point $(a, b, c) \in \mathcal{C}\left(\mathcal{O}_{F_j^\tau}\right)$, with c a unit (since it does not reduce to 0 in the residue field). Dividing through by c^2 then yields $(a/c)^2 - d_j(b/c)^2 = m$ as desired. Pick now a fixed non-square $m \in \mathbb{Z}_\ell^\times$ and for each $j = 2r+1, \ldots, s$ fix a representation $m = a_j^2 - d_j b_j^2$. Take furthermore $x_{2i-1} = 1, x_{2i} = m$ for $i = 1, \ldots, r$.

The corresponding point $x = ((x_i)_{i=1,\ldots,2r}, (x_j)_{j=2r+1,\ldots,s}, m)$ of $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ has the property that $x\tau(x) = m$ is not a square in $\mathbb{Z}_\ell$, and therefore the image of x in $\mathrm{MT}(A)(\mathbb{Z}/\ell^n\mathbb{Z})$ has again the property that $x\tau(x) = [m] \in (\mathbb{Z}/\ell^n\mathbb{Z})^\times$ is not a square. Combined with lemma 5.6.5, this shows that Ψ is not surjective in this case and concludes the proof of theorem 5.6.1 for $\ell \neq 2$.

Notice now that for $\ell = 2$ the lower bound of theorem 5.6.1 is trivial for $n \leq 2$, so we can assume $n \geq 3$. We then remark that (by equation (5.16)) $\mathrm{Hg}(A)(\mathbb{Z}/2^n\mathbb{Z})$ has order at least

$$\prod_{i=1}^{r} 2^{(n-1)[F_{2i-1}:\mathbb{Q}_\ell]} \times \prod_{i=2r+1}^{s} \left| \frac{C^{(i)}(e_i+1)}{C^{(i)}(ne_i)} \right|,$$

which (by the same argument as above, using the second part of lemma 5.6.2) in turn is at least

$$\prod_{i=1}^{r} 2^{(n-1)[F_{2i-1}:\mathbb{Q}_\ell]} \times \prod_{i=2r+1}^{s} \left(2^{f_i}\right)^{(n-1)e_i-1} \geq 2^{g(n-2)}.$$

Furthermore, taking into account the factor coming from the homotheties – namely $(\mathbb{Z}/2^n\mathbb{Z})^\times$ – we find $|\,\mathrm{MT}(A)(\mathbb{Z}/2^n\mathbb{Z})| \geq 2^{(g+1)(n-2)-1}$. Finally, the upper bound for $\ell = 2$ follows trivially from the previous computations and from the second halves of lemmas 5.6.3 and 5.6.5.

5.6.3 Elliptic curves

When the CM abelian variety under consideration is an elliptic curve we can give a complete description of the full adelic Galois representation:

Theorem 5.6.6. *Let A/K be an elliptic curve such that $\mathrm{End}_{\overline{K}}(A)$ is an order in the imaginary quadratic field E. Denote by $\rho_\infty : \mathrm{Gal}(\overline{K}/K) \to \prod_\ell \mathrm{Aut}\, T_\ell A$ the natural adelic representation attached to A, and let G_∞ be its image. For every prime ℓ denote by C_ℓ the group $(\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times$, considered as a subgroup of $\mathrm{Aut}_{\mathbb{Z}_\ell}(\mathcal{O}_E \otimes \mathbb{Z}_\ell) \cong \mathrm{GL}_2(\mathbb{Z}_\ell) \cong \mathrm{Aut}\, T_\ell A$, and let $N(C_\ell)$ be the normalizer of C_ℓ in $\mathrm{GL}_2(\mathbb{Z}_\ell)$.*

1. *Suppose that $E \subseteq K$: then G_∞ is contained in $\prod_\ell C_\ell$, and the index $[\prod_\ell C_\ell : G_\infty]$ does not exceed $3[K:\mathbb{Q}]$. Moreover, the equality $G_{\ell^\infty} = C_\ell$ holds for every prime ℓ unramified in K and such that A has good reduction at all places of K of characteristic ℓ.*

2. *Suppose that $E \not\subseteq K$: then G_∞ is contained in $\prod_\ell N(C_\ell)$ but not in $\prod_\ell C_\ell$, and the index $[\prod_\ell N(C_\ell) : G_\infty]$ is not finite. The intersection $H_\infty = G_\infty \cap \prod_\ell C_\ell$ has index 2 in G_∞, and the index $[\prod_\ell C_\ell : H_\infty]$ does not exceed $6[K:\mathbb{Q}]$. Moreover, the equality $G_{\ell^\infty} = N(C_\ell)$ holds for every prime ℓ unramified in $K \cdot E$ and such that A has good reduction at all places of K of characteristic ℓ.*

Finally, the constants 3 and 6 appearing in parts (1) and (2) respectively can be replaced by 1 and 2 if we further assume that the j-invariant of A is neither 0 nor 1728.

We start by recording the following consequence of theorem 5.5.5:

Corollary 5.6.7. *Let A/K be an elliptic curve admitting complex multiplication (over K) by the imaginary quadratic field E. The group G_{ℓ^∞} is contained in $\mathrm{MT}(A)(\mathbb{Z}_\ell) = C_\ell$, and if A has good reduction at all places of K of characteristic ℓ the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : G_{\ell^\infty}]$ is at most $\frac{1}{2}[K:\mathbb{Q}]$. If in addition ℓ is also unramified in K we have $G_{\ell^\infty} = (\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times$.*

Proof. Since E is quadratic, E and E^* coincide and the reflex norm is simply the identity $T_E \to T_E$, hence $\mathrm{MT}(A) = T_E$ and (in the notation of theorem 5.5.5) F is the trivial group. In particular

$T_E(\mathbb{Z}_\ell) = (\mathcal{O}_E \otimes \mathbb{Z}_\ell)^\times = C_\ell$ contains G_{ℓ^∞} by [124, Corollary 2 to Theorem 5] (cf. also [116, Corollaire on p. 302]): the claim on the index then follows from theorem 5.5.5 upon noticing that $[K : E^*] = [K : E] = \frac{1}{2}[K : \mathbb{Q}]$. Furthermore, if ℓ is unramified in K, then it is also unramified in E, and the remaining assertion $G_{\ell^\infty} = C_\ell = \mathrm{MT}(A)(\mathbb{Z}_\ell)$ follows from part (3) of theorem 5.5.5. □

We shall also need some results concerning elliptic curves A/K that admit complex multiplication over $\overline{K}$ but not over K. We start with the following easy properties of $N(C_\ell)$:

Lemma 5.6.8. *C_ℓ is of index 2 in $N(C_\ell)$. In particular, $N(C_\ell)$ is generated by C_ℓ and any element in $N(C_\ell) \setminus C_\ell$. Furthermore, if H_ℓ is an open subgroup of C_ℓ, then the normalizer of H_ℓ in $\mathrm{GL}_2(\mathbb{Z}_\ell)$ is contained in $N(C_\ell)$.*

Proof. Fix $\omega \in \mathcal{O}_E$ such that $(1, \omega)$ is a $\mathbb{Z}$-basis of $\mathcal{O}_E$. There exist $c, d \in \mathbb{Z}$ such that ω satisfies the quadratic relation $\omega^2 = c\omega + d$. In the $\mathbb{Z}_\ell$-basis $(1, \omega)$ of $\mathcal{O}_E \otimes \mathbb{Z}_\ell$, the group C_ℓ is the subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ given by the invertible matrices that can be written as $\begin{pmatrix} a & bd \\ b & a + bc \end{pmatrix}$ for some $a, b \in \mathbb{Z}_\ell$. We thus see that C_ℓ is given by the intersection of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ with a 2-dimensional plane Π (that defined by the equations $x_{11} + cx_{21} = x_{22}, x_{12} = dx_{21}$, where x_{ij} is the coefficient on the i-th row and j-th column). In particular, for an element $g \in \mathrm{GL}_2(\mathbb{Z}_\ell)$ the condition of normalizing C_ℓ is equivalent to that of stabilizing Π. The latter is a Zariski-closed condition, and since any subgroup H_ℓ of C_ℓ open in the ℓ-adic topology is Zariski-dense in Π we see that if g normalizes H_ℓ, then it stabilizes Π and hence it normalizes C_ℓ. Finally, with the explicit description at hand it is immediate to see that $[N(C_\ell) : C_\ell] = 2$, and that a nontrivial element of $N(C_\ell) \setminus C_\ell$ is given by $\begin{pmatrix} 1 & c \\ 0 & -1 \end{pmatrix}$. □

Lemma 5.6.9. *Suppose A/K is an elliptic curve such that $\mathrm{End}_K(A) = \mathbb{Z}$ but $\mathrm{End}_{\overline{K}}(A)$ is an order in an imaginary quadratic field E: then for every prime ℓ the group G_{ℓ^∞} is contained in $N(C_\ell)$.*

Proof. The field $K^1 = K \cdot E$ is a quadratic extension of K over which all the endomorphisms of A are defined, and the group $G^1_{\ell^\infty} = \rho_{\ell^\infty}\left(\mathrm{Gal}(\overline{K^1}/K^1)\right)$ is a closed subgroup of G_{ℓ^∞} of index at most 2 (hence in particular it is normal and open in G_{ℓ^∞}). Let $R = \mathrm{End}_{\overline{K}}(A)$. Since A admits complex multiplication by R over K^1, we know by [116, §4.5, Corollaire] that $G^1_{\ell^\infty}$ is of finite index in $(R \otimes \mathbb{Z}_\ell)^\times$, which in turn is of finite index in C_ℓ. Thus the normalizer of $G^1_{\ell^\infty}$ is included in $N(C_\ell)$ by lemma 5.6.8, and since $G^1_{\ell^\infty}$ is normal in G_{ℓ^∞} we have $G_{\ell^\infty} \subseteq N(G^1_{\ell^\infty}) \subseteq N(C_\ell)$ as claimed. □

Lemma 5.6.10. *In the situation of the previous lemma, for all primes ℓ the group G_{ℓ^∞} has nonempty intersection with $N(C_\ell) \setminus C_\ell$.*

Proof. For all primes ℓ we have $G_{\ell^\infty} \subseteq N(C_\ell)$. On the other hand, we know by Faltings' theorem that the centralizer of G_{ℓ^∞} in $\mathrm{End}\,(T_\ell A) \otimes \mathbb{Q}_\ell$ equals $\mathrm{End}_K(A) \otimes \mathbb{Q}_\ell = \mathbb{Q}_\ell$. It follows that G_{ℓ^∞} cannot be abelian, for otherwise its centralizer would contain all of G_{ℓ^∞} (which is not contained in the homotheties $\mathbb{Q}_\ell$): in particular, G_{ℓ^∞} must have nonempty intersection with $N(C_\ell) \setminus C_\ell$. □

We can now prove theorem 5.6.6:

Proof. (of theorem 5.6.6) The proof is quite similar to that of theorem 5.5.5, the main differences being that we need to treat all places at the same time and that the action of E needs not be defined over K. Consider first case (1). The inclusion $G_{\ell^\infty} \subseteq C_\ell$ is part of corollary 5.6.7, and implies $G_\infty \subseteq \prod_\ell G_{\ell^\infty} \subseteq \prod_\ell C_\ell$. In particular, G_∞ is abelian, so class field theory allows us to interpret ρ_∞ as a map

$$I_K \xrightarrow{\rho_\infty} \prod_\ell C_\ell$$

that is trivial on K^*. As in the proof of theorem 5.5.5, since we are looking for a *lower* bound on G_∞ no harm is done in replacing I_K by the group of idèles of the Hilbert class field of K; concretely, this means considering the restriction of ρ_∞ to $\prod_{v\in\Omega_K} \mathcal{O}_{K,v}^\times$, where Ω_K is the set of finite places of K. Recall from theorem 5.5.1 that the action of ρ_∞ on a finite idèle $a = (a_v)_{v\in\Omega_K}$ is given by

$$\rho_\infty(a) = \varepsilon(a)\left(N_{K_\ell/E_\ell}(a_\ell^{-1})\right)_{\ell \text{ prime}}.$$

As in the proof of theorem 5.5.5, if we let $\mu(E)$ be the group of roots of unity in E we know that $\ker\varepsilon$ is a subgroup of $\prod_{v\in\Omega_K}\mathcal{O}_{K,v}^\times$ of index at most $|\mu(E)|$, and since E is a quadratic imaginary field we have $|\mu(E)| \leq 6$. Therefore the image of ρ_∞ has index at most $|\ker\varepsilon| \leq 6$ in the image of the map

$$\begin{array}{rccc} \varphi_\infty : & \prod_{v\in\Omega_K}\mathcal{O}_v^\times & \to & \prod_\ell(\mathcal{O}_E\otimes\mathbb{Z}_\ell)^\times = \prod_\ell C_\ell \\ & (a)_v & \mapsto & \left(N_{K_\ell/E_\ell}(a_\ell)\right)_\ell \end{array}$$

given by taking local norms from K_ℓ to E_ℓ. Hence in particular we have

$$\left[\prod_\ell C_\ell : G_\infty\right] \leq 6\left[\prod_\ell C_\ell : \operatorname{Im}\varphi_\infty\right],$$

and it suffices to show that

$$\left[\prod_\ell C_\ell : \operatorname{Im}\varphi_\infty\right] \leq [K:E] = \frac{1}{2}[K:\mathbb{Q}],$$

which follows from [4, Theorem 7 on p. 161] (the global field counterpart of theorem 5.5.3). The remaining assertion of part (1) is exactly the content of corollary 5.6.7.

As for part (2), we have seen in lemmas 5.6.9 and 5.6.10 that in this case G_{ℓ^∞} is contained in $N(C_\ell)$, but not in C_ℓ. If we let $K^1 = K\cdot E$, then A admits complex multiplication by E over K^1, so $\rho_\infty\left(\mathrm{Gal}(\overline{K^1}/K^1)\right)$ is contained in $\prod_\ell C_\ell$ by part (1). Since $\mathrm{Gal}(\overline{K^1}/K^1)$ has index 2 in $\mathrm{Gal}(\overline{K}/K)$ we must have $H_\infty = \rho_\infty\left(\mathrm{Gal}(\overline{K^1}/K^1)\right)$, so that the index $[G_\infty : H_\infty]$ is indeed 2 and applying part (1) we find $[\prod_\ell C_\ell : H_\infty] \leq 3[K^1:\mathbb{Q}] = 6[K:\mathbb{Q}]$; moreover, the index $[\prod_\ell N(C_\ell) : G_\infty]$ is not finite since the same is clearly true for the index $[\prod_\ell N(C_\ell) : \prod_\ell C_\ell]$. Finally, if ℓ is unramified in K^1 we see from corollary 5.6.7 (applied to A/K^1) that G_{ℓ^∞} contains all of C_ℓ, and by lemma 5.6.10 we know that G_{ℓ^∞} also contains an element of $N(C_\ell)\setminus C_\ell$. The equality $G_{\ell^\infty} = N(C_\ell)$ then follows from lemma 5.6.8.

As for the last assertion, notice that if we exclude elliptic curves with j-invariant equal to 0 or 1728 the field of complex multiplication E is neither $\mathbb{Q}(i)$ nor $\mathbb{Q}(\zeta_3)$, so the only roots of unity in E are ± 1. This implies that $\ker\varepsilon$ has index at most 2 in $\prod_{v\in\Omega_K}\mathcal{O}_v^\times$, and the same argument as above shows that the constants 3 and 6 can indeed be replaced by 1 and 2. □

Remark 5.6.11. The following simple example shows that the constants 3 and 6 appearing in the statement of theorem 5.6.6 cannot be improved in general. We consider the elliptic curve A over the field $K = \mathbb{Q}(\zeta_3)$ given by the Weierstrass equation $y^2 = x^3 + 1$. As it is clear, A has complex multiplication (over K) by the full ring of integers of $E = K$. Moreover, all the 2-torsion points of A are defined over K, so G_2 has trivial reduction modulo 2. Hence G_2 is a subgroup of $\ker(\mathbb{Z}_2[\zeta_3]^\times \to \mathbb{F}_2[\zeta_3]^\times)$, and its index in $(\mathcal{O}_E \otimes \mathbb{Z}_2)^\times \cong \mathbb{Z}_2[\zeta_3]^\times$ is divisible by $|\mathbb{F}_2[\zeta_3]^\times| = 3$. Likewise, the fact that the 3-torsion point with coordinates $(0, 1)$ is defined over K shows that the index of G_3 in $(\mathcal{O}_E \otimes \mathbb{Z}_3)^\times$ is divisible by 2. Thus we conclude that the index of G_∞ in $\prod_\ell C_\ell$ is at least $6 = 3[K : \mathbb{Q}]$, so that the constant 3 is indeed sharp. Finally, considering the $\mathbb{Q}$-elliptic curve given by the same Weierstrass equation shows the optimality of part (2): in this case H_∞ is exactly the image of the Galois representation attached to A/K, so we have $[\prod_\ell C_\ell : H_\infty] = 6$ by what we just showed.

5.6.4 Abelian surfaces

An easy direct computation shows that when $\dim A = 2$ the kernel of the reflex norm is always connected, and therefore the group F of theorem 5.5.5 is trivial. Since furthermore simple CM types are automatically non-degenerate in dimension 2, combining theorems 5.5.5 and 5.6.1 we deduce:

Corollary 5.6.12. *Let A/K be an absolutely simple abelian variety of dimension 2. Suppose that A has CM over K by the field E and let ℓ be a prime number such that A has good reduction at all places of K of characteristic ℓ. The group $G_{\ell^\infty} \cap \mathrm{MT}(A)(\mathbb{Z}_\ell)$ has index at most $[K : E^*]$ in $\mathrm{MT}(A)(\mathbb{Z}_\ell)$, hence we have $[K(A[\ell^n]) : K] \geq \frac{1}{[K : E^*]}(1 - 1/\ell)^3 \ell^{3n}$ for $\ell \neq 2$, while for $\ell = 2$ we have $[K(A[2^n]) : K] \geq \frac{1}{2^7[K : E^*]} 2^{3n}$. Finally, if ℓ is unramified in $K \cdot E$ we even have $[K(A[\ell^n]) : K] \geq (1 - 1/\ell)^3 \ell^{3n}$.*

5.7 A family of varieties with small 2-torsion fields

Let $p \geq 3$ be a prime number and K_p be the cyclotomic field $\mathbb{Q}(\zeta_p)$. We let C_p be the unique smooth K_p-curve birational to $y^p = x(1 - x)$ and $J(p)$ be its Jacobian, again over K_p. It is clear that C_p admits an action of μ_p, so $J(p)$ is a CM abelian variety, admitting complex multiplication over K_p by the full ring of integers of K_p. Notice furthermore that C_p is birational to the curve

$$z^2 = w^p + 1/4$$

(just set $x = z + 1/2$, $y = -w$), so it is hyperelliptic of genus $\frac{p-1}{2}$. Direct inspection of the model $y^p = x(1 - x)$ reveals that C_p is smooth away from p, so $J(p)$ has everywhere good reduction over K_p except perhaps at the unique place dividing p. The reflex field is $K_p^* = K_p$. Let us compute the CM type S of $J(p)$: in the basis $\omega_j := w^j \frac{dw}{z}$ $(j = 0, \ldots, \frac{p-3}{2})$ of the space of differentials on C_p, the action of ζ_p is given by $[\zeta_p]^* \omega_j = \zeta_p^{j+1} \omega_j$, hence the CM type, considered as a subset of $\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times$, is $\left\{1, \ldots, \frac{p-1}{2}\right\}$. Equivalently,

$$S = \left\{ g \in \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times \;\middle|\; 2\langle g \rangle < p \right\},$$

where $\langle g \rangle$ is the unique integer lying in the interval $[0, p-1]$ that is congruent to g modulo p. This description shows that our CM type agrees with the type S_1 described in [66], which by [66, Lemma 1] is nondegenerate (cf. also [52]): thus we have $\operatorname{rank} \mathrm{MT}(A) = \dim A + 1 = \frac{p+1}{2}$.

Let now $\beta_1, \dots, \beta_p$ be the roots of $w^p + 1/4 = 0$ in $\overline{\mathbb{Q}}$, and let $P_i = (\beta_i, 0)$ be the corresponding points of C_p (in the coordinates (w, z)). Finally, for $i = 1, \dots, p$ let d_i denote the divisor $(P_i) - (\infty)$. It is known (see for example [10, §5.1]) that the 2-torsion of $J(p)$ is an $\mathbb{F}_2$-vector space of dimension $p-1$ spanned by the d_i's, which are only subject to the linear relation $\sum_{i=1}^{p} d_i = [0]$. It follows that the 2-torsion field $K_p(J(p)[2]) = K_p(\{\beta_i\}) = K_p\left(\sqrt[p]{1/4}\right)$ has degree p over K_p, so for $\ell = 2$ and $n = 1$ the ratio $\ell^{n \operatorname{rank} \mathrm{MT}(A)} \,/\, [K(A[\ell^n]) : K]$ is given by

$$\frac{2^{\operatorname{rank} \mathrm{MT}(A)}}{[K(J(p)[2]) : K]} = \frac{2^{(p+1)/2}}{p} = \frac{2^{\dim J(p)+1}}{2 \dim J(p) + 1},$$

which shows in particular that, as claimed in the introduction, the optimal bound on the quantity $\ell^{n \operatorname{rank} \mathrm{MT}(A)} \,/\, [K(A[\ell^n]) : K]$ grows at least exponentially in the dimension of A.

Chapter 6

On the ℓ-adic Galois representations attached to nonsimple abelian varieties

6.1 Introduction

Let K be a field finitely generated over its prime subfield, and let A be an abelian variety over K. The action of the absolute Galois group of K on the various Tate modules $T_\ell A$ (for $\ell \neq \operatorname{char} K$) gives a (compatible) family of ℓ-adic representations of the absolute Galois group of K, and most of the relevant information is encoded neatly in a certain family of algebraic groups (denoted $H_\ell(A)$ in what follows, cf. definitions 6.2.5 and 6.5.5). It is thus very natural to try and understand the Galois action on nonsimple varieties in terms of the groups H_ℓ; the main results of this chapter are several sufficient criteria for the equality $H_\ell(A \times B) \cong H_\ell(A) \times H_\ell(B)$ to hold. We start by discussing the case $\operatorname{char} K = 0$, which is technically simpler, and prove for example the following ℓ-adic version, and mild generalization, of a Hodge-theoretical result of Hazama [36]:

Theorem 6.4.1. *Let K be a finitely generated field of characteristic zero, A_1 and A_2 be K-abelian varieties, and ℓ be a prime number. For $i = 1, 2$ let $\mathfrak{h}_i$ be the Lie algebra of $H_\ell(A_i)$. Suppose that the following hold:*

1. *for $i = 1, 2$, the algebra $\mathfrak{h}_i$ is semisimple, so that we can write $\mathfrak{h}_i \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{h}_{i,1} \oplus \cdots \oplus \mathfrak{h}_{i,n_i}$, where every $\mathfrak{h}_{i,j}$ is simple;*

2. *for $i = 1, 2$, there exists a decomposition $V_\ell(A_i) \otimes \overline{\mathbb{Q}_\ell} \cong V_{i,1} \oplus \cdots \oplus V_{i,n_i}$ such that the action of $\mathfrak{h}_i \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{h}_{i,1} \oplus \cdots \oplus \mathfrak{h}_{i,n_i}$ on $V_{i,1} \oplus \cdots \oplus V_{i,n_i}$ is componentwise and $\mathfrak{h}_{i,j}$ acts faithfully on $V_{i,j}$;*

3. *for all distinct pairs (i, j) and (i', j') for which there exists an isomorphism $\varphi : \mathfrak{h}_{i,j} \to \mathfrak{h}_{i',j'}$ there is an irreducible $\mathfrak{h}_{i,j}$-representation W such that all simple $\mathfrak{h}_{i,j}$-submodules of $V_{i,j}$ and of $\varphi^*\left(V_{i',j'}\right)$ are isomorphic to W, and the highest weight defining W is stable under all automorphisms of $\mathfrak{h}_{i,j}$.*

Then either $\mathrm{Hom}_{\overline{K}}(A_1, A_2) \neq 0$ *or* $H_\ell(A_1 \times A_2) \cong H_\ell(A_1) \times H_\ell(A_2)$.

From this theorem we deduce many easily applicable criteria, including for example the following result on low-dimensional abelian varieties.

Corollary 6.4.5. *Let K be a finitely generated subfield of $\mathbb{C}$ and $A_1, \ldots, A_n$ be absolutely simple K-abelian varieties of dimension at most 2, pairwise non-isogenous over $\overline{K}$. Let $k_1, \ldots, k_n$ be positive integers and A be a K-abelian variety that is $\overline{K}$-isogenous to $\prod_{i=1}^n A_i^{k_i}$. Then we have $H_\ell(A) \cong \prod_{i=1}^n H_\ell(A_i)$, and the Mumford-Tate conjecture holds for A.*

On the other hand, as the conditions in theorem 6.4.1 are often not easy to check, it would be desirable to describe families of abelian varieties for which they are known to hold; in this direction we prove a result inspired by a paper of Ichikawa [44], where a sufficient criterion is given for the equality $H(A \times B) \cong H(A) \times H(B)$ to hold for the Hodge groups of complex abelian varieties. The criterion is expressed in terms of the *relative dimensions* of the factors:

Definition 6.1.1. Let K be any field and A be an absolutely simple K-abelian variety, so that $\mathrm{End}^0_{\overline{K}}(A) = \mathrm{End}_{\overline{K}}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a division algebra, with center a number field E (either totally real or CM). The degree of $\mathrm{End}^0_{\overline{K}}(A)$ over E is a perfect square, which we write as d^2; by *type* of A we mean the type of $\mathrm{End}^0_{\overline{K}}(A)$ in the Albert classification. The relative dimension of A is then given by

$$\mathrm{reldim}(A) = \begin{cases} \dfrac{\dim A}{de}, & \text{if } A \text{ is of type I, II or III} \\ \dfrac{2\dim A}{de}, & \text{if } A \text{ is of type IV} \end{cases}$$

Note that $d = 1$ if A is of type I, and $d = 2$ if A is of type II or III.

A Ribet-style lemma (proved in section 6.3) that slightly generalizes results found in the literature, combined with techniques due to Pink [98] and Larsen-Pink [57], allows us to prove the following ℓ-adic analogue of Ichikawa's theorem, which has exactly the same form as the corresponding Hodge-theoretical result:

Theorem 6.4.7. *Let K be a finitely generated field of characteristic zero and A'_i, A''_j (for $i = 1, \ldots, n$ and $j = 1, \ldots, m$) be absolutely simple K-abelian varieties of odd relative dimension that are pairwise non-isogenous over $\overline{K}$. Suppose every A'_i is of type I, II or III in the sense of Albert, and every A''_j is of type IV. Let A be a K-abelian variety that is $\overline{K}$-isogenous to $\prod_{i=1}^n A'_i \times \prod_{j=1}^m A''_j$: then*

$$H_\ell(A) \cong \prod_{i=1}^n H_\ell\left(A'_i\right) \times H_\ell\left(\prod_{j=1}^m A''_j\right).$$

In section 6.5 we then discuss to which extent the previous results apply to finitely generated fields of positive characteristic. It turns out that in this setting the most natural definition of $H_\ell(A)$ is different, and that some additional technical hypotheses must be added to our main results. Theorems 6.5.7 and 6.5.9 are positive-characteristic versions of theorems 6.4.1 and 6.4.7 respectively; they are slightly weaker than their characteristic-zero counterparts, but are still qualitatively very similar.

Finally, in section 6.6 we apply our results to nonsimple varieties of dimension at most 5 defined over finitely generated subfields of $\mathbb{C}$; by studying the product structure of H_ℓ we prove the Mumford-Tate conjecture for most such varieties, and in all cases we are able to reproduce in the arithmetical setting results obtained in [82] for their Hodge group. Note that [82] makes ample use of compactness arguments (for real semisimple groups) that are not available in the ℓ-adic context and thus need to be replaced in our setting.

6.2 Preliminaries

6.2.1 Notation

Throughout the chapter the letter A will be reserved for an abelian variety defined over a field K, which we suppose to be finitely generated (over its prime subfield). A field K will be said to be a "finitely generated subfield of $\mathbb{C}$" if it is finitely generated over $\mathbb{Q}$ and a distinguished embedding $\sigma : K \hookrightarrow \mathbb{C}$ has been fixed. If A is an abelian variety defined over a finitely generated subfield of $\mathbb{C}$, we will write $A_{\mathbb{C}}$ for the base-change of A to $\mathbb{C}$ along σ; the symbol $V(A)$ will then denote the first homology group $H_1\left(A_{\mathbb{C}}(\mathbb{C}), \mathbb{Q}\right)$. We will also denote ℓ a prime number, and write $V_\ell(A)$ for $T_\ell(A) \otimes \mathbb{Q}_\ell$, where $T_\ell(A)$ is as usual the ℓ-adic Tate module of A.

If G is an algebraic group we shall write G^{der} for its derived subgroup, $Z(G)$ for the connected component of its center, and G^0 for the connected component of the identity; when $\mathfrak{h}$ is a reductive Lie algebra we shall write $\mathfrak{h}^{\mathrm{ss}}$ for its semisimple part. Finally, if $\varphi : \mathfrak{g} \to \mathfrak{h}$ is a morphism of Lie algebras and $\rho : \mathfrak{h} \to \mathfrak{gl}(V)$ is a representation of $\mathfrak{h}$, we denote $\varphi^*(V)$ the representation $\rho \circ \varphi$ of $\mathfrak{g}$.

Definition 6.2.1. When $\mathfrak{h}$ is a classical Lie algebra (i.e. of Lie type A_l, B_l, C_l, or D_l), we call **standard representation** of $\mathfrak{h}$ the one coming from the defining representation of the corresponding algebraic group. It is in all cases the representation with highest weight ω_1 (in the notation of Bourbaki [18, Planches I-IV]).

6.2.2 The Hodge group

We now briefly recall the notion of Hodge group of an abelian variety (defined over an arbitrary subfield F of $\mathbb{C}$), referring the reader to [79] for more details. To stress that F need not be finitely generated, we depart from our standard notation A and denote X an abelian variety defined over F; we denote $X_{\mathbb{C}}$ the base-change of X to $\mathbb{C}$. The $\mathbb{Q}$-vector space $V(X) = H_1\left(X_{\mathbb{C}}(\mathbb{C}), \mathbb{Q}\right)$ is naturally endowed with a Hodge structure of type $(-1,0) \oplus (0,-1)$, that is, a decomposition of $\mathbb{C}$-vector spaces $V(X) \otimes \mathbb{C} \cong V(X)^{-1,0} \oplus V(X)^{0,-1}$ such that $\overline{V(X)^{-1,0}} = V(X)^{0,-1}$.
Let $\mu_\infty : \mathbb{G}_{m,\mathbb{C}} \to \mathrm{GL}\left(V(X)_{\mathbb{C}}\right)$ be the unique cocharacter such that $z \in \mathbb{C}^*$ acts as multiplication by z on $V(X)^{-1,0}$ and trivially on $V(X)^{0,-1}$. The **Mumford-Tate** group of X is the $\mathbb{Q}$-Zariski closure of the image of μ_∞, that is to say the smallest $\mathbb{Q}$-algebraic subgroup $\mathrm{MT}(X)$ of $\mathrm{GL}(V(X))$ such that μ_∞ factors through $\mathrm{MT}(X)_{\mathbb{C}}$. It is not hard to show that $\mathrm{MT}(X)$ contains the torus of homotheties in $\mathrm{GL}(V(X))$.

Definition 6.2.2. The **Hodge group** of X is $H(X) = \left(\mathrm{MT}(X) \cap \mathrm{SL}(V(X))\right)^0$.

Remark 6.2.3. The group $\mathrm{MT}(X)$ can be recovered from the knowledge of $H(X)$: indeed, $\mathrm{MT}(X)$ is the almost-direct product of $\mathbb{G}_m$ and $H(X)$ inside $\mathrm{GL}(V(X))$, where $\mathbb{G}_m$ is the central torus of homotheties.

It is well known that the group $H(X)$ is connected and reductive, and that there is an isomorphism $\mathrm{End}^0_{\overline{F}}(X) \cong \mathrm{End}(V(X))^{H(X)}$. Moreover, if λ is a polarization of $X_{\mathbb{C}}$ and φ is the bilinear form induced on $V(X)$ by λ, the group $H(X)$ is contained in $\mathrm{Sp}(V(X), \varphi)$. It is also easy to show that when the F-abelian varieties X_1 and X_2 are isogenous over $\mathbb{C}$ the groups $H(X_1)$ and $H(X_2)$ are isomorphic, and that when $X_{\mathbb{C}}$ has no simple factor of type IV the group $H(X)$ is semisimple. Finally, we also have some information on the behaviour of $H(X)$ with respect to products:

Proposition 6.2.4. *Let F be a subfield of $\mathbb{C}$ and X_1, X_2 be abelian varieties defined over F. The group $H(X_1 \times X_2)$ is contained in $H(X_1) \times H(X_2)$, and it projects surjectively on both factors.*
Let $X_1, \dots, X_k$ be absolutely simple F-abelian varieties that are pairwise non-isogenous over $\mathbb{C}$, and let $n_1, \dots, n_k$ be positive integers. The groups $H(X_1^{n_1} \times \cdots \times X_k^{n_k})$ and $H(X_1 \times \cdots \times X_k)$ are isomorphic.

6.2.3 The groups $H_\ell(A)$

Let now K be a finitely generated field of characteristic zero, A be an abelian variety defined over K, and ℓ be a prime number; recall that we set $V_\ell(A) = T_\ell(A) \otimes \mathbb{Q}_\ell$. The action of $\mathrm{Gal}\left(\overline{K}/K\right)$ on the torsion points of A induces a representation $\rho_\ell : \mathrm{Gal}\left(\overline{K}/K\right) \to \mathrm{GL}(V_\ell(A)) \cong \mathrm{GL}_{2\dim A}(\mathbb{Q}_\ell)$; the Zariski closure of the image of ρ_ℓ is called the **algebraic monodromy group at** ℓ, and is denoted $G_\ell(A)$. As in the Hodge-theoretical case, it is known that $G_\ell(A)$ contains the homotheties (Bogomolov [13]), so that $G_\ell(A)$ is determined by its intersection with $\mathrm{SL}(V_\ell(A))$. This intersection is our main object of study.

Definition 6.2.5. Let K be a finitely generated field of characteristic zero and A be a K-abelian variety. We set $H_\ell(A) = (G_\ell(A) \cap \mathrm{SL}(V_\ell(A)))^0$.

Suppose now that we have fixed an embedding $K \hookrightarrow \mathbb{C}$, so that we can speak of the Hodge group of A. The Mumford-Tate conjecture predicts that the group $H_\ell(A)$ should be an ℓ-adic analogue of $H(A)$, and the two groups are indeed known to share many important properties. It is clear by definition that $H_\ell(A)$ is connected; furthermore, by the comparison isomorphism of étale cohomology we can write $V_\ell(A) \cong V(A) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$, and since $V(A)$ is equipped with a bilinear form φ (induced by a polarization) we obtain by extension of scalars a bilinear form φ_ℓ on $V_\ell(A)$. It is then possible to show that the inclusion $H_\ell(A) \subseteq \mathrm{Sp}(V_\ell(A), \varphi_\ell)$ holds.
Deeper properties of $H_\ell(A)$ are intimately related to Tate's conjecture for abelian varieties, and we summarize them in the following theorem:

Theorem 6.2.6. *(Faltings [26], [27]) Let K be a finitely generated field of characteristic zero, ℓ be a prime number, and A, B be K-abelian varieties. Then $G_\ell(A)$ is a reductive group, and we have*

$$\mathrm{Hom}_{\mathbb{Q}_\ell[G_\ell(A\times B)]}\left(V_\ell(A), V_\ell(B)\right) \cong \mathrm{Hom}_K(A, B) \otimes \mathbb{Q}_\ell.$$

In particular we have $\mathrm{End}(V_\ell(A))^{G_\ell(A)} \cong \mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}_\ell$.

Corollary 6.2.7. *Let K be a finitely generated field of characteristic zero, A and B be abelian varieties defined over K, ℓ be a prime number, and $\mathfrak{h}_\ell$ be the Lie algebra of $H_\ell(A \times B)$. Suppose* $\mathrm{Hom}_{\mathfrak{h}_\ell}(V_\ell(A), V_\ell(B)) \neq 0$*: then* $\mathrm{Hom}_{\overline{K}}(A, B) \neq 0$.

Proof. There is a finite extension K' of K such that the Zariski closure G_ℓ of the image of the representation $\mathrm{Gal}\left(\overline{K'}/K'\right) \to \mathrm{Aut}\left(V_\ell(A \times B)\right)$ is connected. We want to show that $\mathrm{Hom}_{K'}(A, B) \neq 0$. By the previous theorem it is enough to prove that $\mathrm{Hom}_{\mathbb{Q}_\ell[G_\ell]}(V_\ell(A), V_\ell(B))$ is nontrivial. As G_ℓ is connected, an element of $\mathrm{Hom}(V_\ell(A), V_\ell(B))$ is G_ℓ-equivariant if and only if it is equivariant for the action of the Lie algebra $\mathfrak{g}_\ell$ of G_ℓ. On the other hand, we know there is an isomorphism $\mathfrak{g}_\ell \cong \mathfrak{h}_\ell \oplus \mathbb{Q}_\ell$, where the factor $\mathbb{Q}_\ell$ corresponds to the homotheties. Since any linear map commutes with the action of the homotheties we have $\mathrm{Hom}_{\mathbb{Q}_\ell[G_\ell]}(V_\ell(A_1), V_\ell(A_2)) \cong \mathrm{Hom}_{\mathfrak{h}_\ell}(V_\ell(A_1), V_\ell(A_2))$, and the latter space is nontrivial by hypothesis. Thus $\mathrm{Hom}_{K'}(A_1, A_2)$, and a fortiori $\mathrm{Hom}_{\overline{K}}(A_1, A_2)$, are both nontrivial. □

Notice furthermore that the group $H_\ell(A)$ is unchanged by finite extensions of the base field K, and that if A, B are K-abelian varieties that are $\overline{K}$-isogenous we have $H_\ell(A) \cong H_\ell(B)$.
Moreover, $H_\ell(A)$ is semisimple when $A_{\overline{K}}$ does not have any simple factor of type IV (the proof of this fact being the same as for Hodge groups, cf. again [79], especially proposition 1.24), and it has the same behaviour as $H(A)$ with respect to products:

Proposition 6.2.8. *Let K be a finitely generated field of characteristic zero and A_1, A_2 be K-abelian varieties. The group $H_\ell(A_1 \times A_2)$ is contained in $H_\ell(A_1) \times H_\ell(A_2)$, and it projects surjectively on both factors.*
Let $A_1, \ldots, A_k$ be absolutely simple K-abelian varieties that are pairwise non-isogenous over $\overline{K}$, and let $n_1, \ldots, n_k$ be positive integers. The groups $H_\ell(A_1^{n_1} \times \cdots \times A_k^{n_k})$ and $H_\ell(A_1 \times \cdots \times A_k)$ are isomorphic.

We also have some information about the structure of $V_\ell(A)$ as a representation of $H_\ell(A)$:

Theorem 6.2.9. *(Pink, [98, Corollary 5.11]) Let K be a finitely generated field of characteristic zero, A be a K-abelian variety, ℓ be a prime number, and $\mathfrak{h}_\ell(A)$ be the Lie algebra of $H_\ell(A)$. Write $\mathfrak{h}_\ell(A) \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{c} \oplus \bigoplus_{i=1}^n \mathfrak{h}_i$, where $\mathfrak{c}$ is abelian and each $\mathfrak{h}_i$ is simple. Let W be a simple submodule of $V_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ for the action of $(\mathfrak{h}_\ell(A) \otimes \overline{\mathbb{Q}_\ell})$, decomposed as $W \cong C \otimes \bigotimes_{i=1}^n W_i$, where each W_i is a simple module over $\mathfrak{h}_i$ and C is a 1-dimensional representation of $\mathfrak{c}$. Then:*

1. *each $\mathfrak{h}_i$ is of classical type (i.e. of Lie type A_l, B_l, C_l or D_l for some l);*
2. *if W_i is nontrivial, then the highest weight of $\mathfrak{h}_i$ in W_i is minuscule.*

Remark 6.2.10. This theorem is stated in [98] only for number fields. The version for finitely generated fields follows easily by a specialization argument (cf. also proposition 6.2.11 below).

For the reader's convenience and future reference, we reproduce the full list of minuscule weights for classical Lie algebras, as given for example in [18] (Chapter 8, Section 3 and Tables 1 and 2); the last column of this table contains $+1$ if the corresponding representation is orthogonal, -1 if it is symplectic, and 0 if it is not self-dual.

Root system	**Minuscule weight**	**Dimension**	**Duality properties**
$A_l\,(l \geq 1)$	$\omega_r, 1 \leq r \leq l$	$\binom{l+1}{r}$	$(-1)^r$, if $r = \dfrac{l+1}{2}$ 0, if $r \neq \dfrac{l+1}{2}$
$B_l\,(l \geq 2)$	ω_l	2^l	$+1$, if $l \equiv 3, 0 \pmod 4$ -1, if $l \equiv 1, 2 \pmod 4$
$C_l\,(l \geq 3)$	ω_1	$2l$	-1
$D_l\,(l \geq 4)$	ω_1	$2l$	$+1$
	ω_{l-1}, ω_l	2^{l-1}	$+1$, if $l \equiv 0 \pmod 4$ -1, if $l \equiv 2 \pmod 4$ 0, if $l \equiv 1 \pmod 2$

TABLE 6.1: Minuscule weights

6.2.4 Known results towards the Mumford-Tate conjecture

Let K be again a field finitely generated over $\mathbb{Q}$, and A be an abelian variety over K. Fix any embedding $\sigma : K \hookrightarrow \mathbb{C}$, so that we can regard K as a subfield of $\mathbb{C}$, and the Mumford-Tate and Hodge groups of A are defined. The celebrated Mumford-Tate conjecture predicts that the equality $G_\ell(A)^0 = \mathrm{MT}(A) \otimes \mathbb{Q}_\ell$ should hold for every prime ℓ; equivalently, for every A and ℓ we should have $H_\ell(A) \cong H(A) \otimes \mathbb{Q}_\ell$. Note that both sides of this equality are invariant under finite extensions of K and isogenies: in particular, if A and B are K-abelian varieties that are $\overline{K}$-isogenous, the conjecture holds for A if and only if it holds for B.
Even though the general case of the conjecture is still wide open, many partial results have proven, and we shall now recall a number of them that we will need in what follows. Let us start with the following proposition, which allows a reduction of the problem to the case of K being a number field:

Proposition 6.2.11. *(Serre, Noot, [91, Proposition 1.3]) Let ℓ be a prime, K be a finitely generated subfield of $\mathbb{C}$ and A be a K-abelian variety. There exist a number field L, a specialization B of A over L, and identifications $H_1(A_{\mathbb{C}}(\mathbb{C}), \mathbb{Q}) \cong H_1(B_{\mathbb{C}}(\mathbb{C}), \mathbb{Q})$ and $T_\ell(A) \cong T_\ell(B)$ (compatible with the comparison isomorphism in étale cohomology) such that $\mathrm{MT}(A) = \mathrm{MT}(B)$ and $G_\ell(A) = G_\ell(B)$ under the given identifications.*

This proposition implies in particular that most results which are known for number fields and depend on a single prime ℓ automatically propagate to finitely generated subfields of $\mathbb{C}$. This applies to all the theorems we list in this section, some of which were originally stated only for number fields.

Theorem 6.2.12. *(Piatetskii-Shapiro, Borovoi, Deligne [23, I, Proposition 6.2]) Let K be a finitely generated subfield of $\mathbb{C}$ and A be a K-abelian variety. For every prime ℓ we have the inclusion $G_\ell(A)^0 \subseteq \mathrm{MT}(A) \otimes \mathbb{Q}_\ell$.*

Theorem 6.2.13. *(Pink, [59, Theorem 4.3]) Let K be a finitely generated subfield of $\mathbb{C}$ and A be a K-abelian variety. Suppose that the equality $\mathrm{rk}(H(A)) = \mathrm{rk}(H_\ell(A))$ holds for one prime ℓ: then*

$H_\ell(A) = H(A) \otimes \mathbb{Q}_\ell$ holds for every prime ℓ. In particular, if the Mumford-Tate conjecture holds for one prime, then it holds for every prime.

Theorem 6.2.14. *(Vasiu, [137, Theorem 1.3.1]; cf. also Ullmo-Yafaev, [135, Corollary 2.11]) Let K be a finitely generated subfield of $\mathbb{C}$ and A be a K-abelian variety. For every prime ℓ we have $Z(H_\ell(A)) \cong Z(H(A)) \otimes \mathbb{Q}_\ell$. In particular, the Mumford-Tate conjecture is true for CM abelian varieties.*

Remark 6.2.15. The CM case of the Mumford-Tate conjecture was first proved by Pohlmann [101].

The following proposition follows immediately upon combining the previous three theorems:

Proposition 6.2.16. *Let K be a finitely generated subfield of $\mathbb{C}$ and A be a K-abelian variety. Suppose that for one prime number ℓ we have $\mathrm{rk}(H(A)^{\mathrm{der}}) \leq \mathrm{rk}(H_\ell(A)^{\mathrm{der}})$: then the Mumford-Tate conjecture holds for A. The same is true if (for some prime ℓ) we have $\mathrm{rk}\, H(A) \leq \mathrm{rk}\, H_\ell(A)$.*

In a different direction, many results are known for absolutely simple abelian varieties of specific dimensions:

Theorem 6.2.17. *(Serre, [119]) The Mumford-Tate conjecture is true for elliptic curves (over finitely generated subfields of $\mathbb{C}$).*

Theorem 6.2.18. *(Tanke'ev, Ribet, [112, Theorems 1, 2 and 3]) The Mumford-Tate conjecture is true for absolutely simple abelian varieties of prime dimension (over finitely generated subfields of $\mathbb{C}$).*

Theorem 6.2.19. *(Moonen, Zarhin, [80]) Let K be a finitely generated subfield of $\mathbb{C}$ and A be an absolutely simple K-abelian variety of dimension 4. If $\mathrm{End}_{\overline{K}}(A) \neq \mathbb{Z}$, then the Mumford-Tate conjecture holds for A. If $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$, then either for all primes ℓ we have $H_\ell(A) \cong \mathrm{Sp}_{8,\mathbb{Q}_\ell}$ and Mumford-Tate holds for A, or else for all ℓ the group $H_\ell(A)$ is a $\mathbb{Q}_\ell$-form of SL_2^3.*

Remark 6.2.20. The preprint [155] announces a proof of the Mumford-Tate conjecture for absolutely simple abelian fourfolds A with $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$. In what follows we shall not need this fact, whose only effect would be to slightly simplify the statement of proposition 6.6.2.

There are some common elements to the proofs of all the dimension-specific results we just listed, and we shall try to capture them in definition 6.2.22 below. We now try to motivate this definition. As the group $H_\ell(A)$ is reductive and connected, most of its structure is encoded by the $\mathbb{Q}_\ell$-Lie algebra $\mathfrak{h}_\ell(A) = \mathrm{Lie}(H_\ell(A))$; extending scalars to $\overline{\mathbb{Q}_\ell}$, the Lie algebra $\mathfrak{h}_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ can be written as $\mathfrak{c} \oplus \bigoplus_{i=1}^n \mathfrak{h}_i$, with $\mathfrak{c}$ abelian and each $\mathfrak{h}_i$ simple. The proofs of theorems 6.2.17 and 6.2.18 yield information about the structure of this Lie algebra:

Proposition 6.2.21. *Let K be a finitely generated subfield of $\mathbb{C}$ and A/K be an absolutely simple abelian variety whose dimension is either 1 or a prime number. Fix a prime ℓ and let $\mathfrak{h}_\ell(A)$ be the Lie algebra of $H_\ell(A)$. Suppose A is not of type IV. Then the following hold:*

- *the Lie algebra $\mathfrak{h}_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ admits a decomposition $\mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$, where each simple factor $\mathfrak{h}_i$ is of Lie type $\mathfrak{sp}_k$ for some k;*

- *for each $i = 1, \dots, n$ there exists a (not necessarily simple) $\mathfrak{h}_i$-module W_i such that $V_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ is isomorphic to $W_1 \oplus \cdots \oplus W_n$, the action of $\mathfrak{h}_1 \oplus \ldots \oplus \mathfrak{h}_n$ on $W_1 \oplus \cdots \oplus W_n$ is componentwise, and $\mathfrak{h}_i$ acts faithfully on W_i;*
- *every module W_i is a direct sum of copies of the standard representation of $\mathfrak{h}_i$ (cf. definition 6.2.1).*

Trying to isolate the essential features of this proposition, and taking into account theorem 6.2.9, we are led to the following definition:

Definition 6.2.22. Let K be a finitely generated field of characteristic zero, A/K be an abelian variety, and $\mathfrak{h}_\ell(A)$ be the Lie algebra of $H_\ell(A)$. We can write $\mathfrak{h}_\ell(A) \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{c} \oplus \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$, where $\mathfrak{c}$ is abelian and each factor $\mathfrak{h}_i$ is simple and (by theorem 6.2.9) of classical type. We say that A is of **general Lefschetz type** if it is absolutely simple, not of type IV, and for every prime ℓ the following hold:

1. for each $i = 1, \dots, n$ there exists a (not necessarily simple) $\mathfrak{h}_i$-module W_i such that $V_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ is isomorphic to $W_1 \oplus \cdots \oplus W_n$, where the action of $\mathfrak{h}_1 \oplus \ldots \oplus \mathfrak{h}_n$ on $W_1 \oplus \cdots \oplus W_n$ is componentwise, and $\mathfrak{h}_i$ acts faithfully on W_i;
2. if the simple Lie algebra $\mathfrak{h}_i$ is of Lie type A_l, the rank l is odd and W_i is a direct sum of copies of $\bigwedge^{\frac{l+1}{2}} \mathrm{Std}$, where Std is the standard representation of $\mathfrak{h}_i$ (cf. definition 6.2.1);
3. if the simple algebra $\mathfrak{h}_i$ is of Lie type B_l, the module W_i is a direct sum of copies of the (spinor) representation defined by the highest weight ω_l (in the notation of [18, Planches I-IV]);
4. if the simple algebra $\mathfrak{h}_i$ is of Lie type C_l or D_l, the module W_i is a direct sum of copies of the standard representation of $\mathfrak{h}_i$.

Remark 6.2.23. As proved in [87, Lemma 2.3], when A is a complex abelian variety of type I or II the action of the Lefschetz group of A on $V(A) \otimes \mathbb{C}$ has precisely this structure.

Several instances of this situation have been studied, for example in a series of papers by Banaszak, Gajda and Krasoń. Among various other results, for abelian varieties of type I and II they prove:

Theorem 6.2.24. *(Theorems 6.9 and 7.12 of [6]) Let K be a finitely generated subfield of $\mathbb{C}$ and A/K be an absolutely simple abelian variety of type I or II. Suppose that $h = \mathrm{reldim}(A)$ is odd: then for every prime ℓ the simple factors of $H_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ are of type Sp_{2h}. Furthermore, the Mumford-Tate conjecture holds for A.*

Remark 6.2.25. It is clear from the proof of [6, Lemma 4.13] that any variety as in the previous statement is of general Lefschetz type. Moreover, the result also holds for $h = 2$: this is not stated explicitly in [6], but follows essentially from the same proof (cf. also [21, Theorem 8.5], which covers the case of abelian fourfolds of relative dimension 2).

Another paper by the same authors, [8], deals with varieties of type III:

Proposition 6.2.26. *Let K be a finitely generated subfield of $\mathbb{C}$ and A/K be an absolutely simple abelian variety of type III. Suppose that $h = \mathrm{reldim}(A)$ is odd: then for every ℓ the simple factors of $(\mathrm{Lie}\, H_\ell(A)) \otimes \overline{\mathbb{Q}_\ell}$ are either of type $\mathfrak{so}_{2h}$ or of type $\mathfrak{sl}_{l+1}$, where $l+1$ is a power of 2. Furthermore, A is of general Lefschetz type.*

Remark 6.2.27. Note that the authors of [8] claim a stronger statement, namely the fact that the simple factors of $H_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ can only be of type SO_{2h} and that, under the same hypotheses, Mumford-Tate holds for A. The proof of [8, Lemma 4.13], however, fails to take into account the minuscule orthogonal representations whose dimension is congruent to 2 modulo 4 (those corresponding to algebras of type $\mathfrak{sl}_{l+1}$ acting on $\Lambda^{\frac{l+1}{2}}\mathrm{Std}$, when $l \geq 3$ and $l+1$ is a power of 2); as a result, the statements of [8, Theorems 4.19 and 5.11] need to be amended as we did in proposition 6.2.26.

6.3 Preliminary lemmas

We now start proving some lemmas on algebraic groups and Lie algebras we will repeatedly need throughout the chapter.

Lemma 6.3.1. *Let $G \hookrightarrow G_1 \times G_2$ be an inclusion of algebraic groups over a field of characteristic zero. Suppose that G, G_1 and G_2 are reductive and connected, and that the projections of G on G_1 and G_2 are surjective. If $\mathrm{rk}\, G$ equals $\mathrm{rk}(G_1) + \mathrm{rk}(G_2)$, then the inclusion is an isomorphism.*

Proof. We show that G is open and closed in $G_1 \times G_2$. It is closed because every algebraic subgroup is, and it is open since G and $G_1 \times G_2$ have the same Lie algebra by [35, Lemma 3.1]. □

Lemma 6.3.2. *Let G be a $\mathbb{Q}$-simple algebraic group. If G is semisimple and the number of simple factors of $G_{\overline{\mathbb{Q}}}$ is at most 3, then there is a set of primes L of positive density such that for every ℓ in L the group $G_{\mathbb{Q}_\ell}$ is simple.*

Proof. Let n be the number of simple factors of $G_{\overline{\mathbb{Q}}}$; if $n = 1$ there is nothing to prove, so we can assume n is 2 or 3. The permutation action of $\mathrm{Gal}\left(\overline{\mathbb{Q}}/\mathbb{Q}\right)$ on the simple factors of $G_{\overline{\mathbb{Q}}}$ determines a map $\rho : \mathrm{Gal}\left(\overline{\mathbb{Q}}/\mathbb{Q}\right) \to S_n$, and the assumption that G is $\mathbb{Q}$-simple implies that the image of ρ is a transitive subgroup of S_n. As $n \leq 3$, we see that the image of ρ contains an n-cycle g. By the Chebotarev density theorem there exists a set of primes L of positive density such that $\rho\left(\mathrm{Gal}\left(\overline{\mathbb{Q}_\ell}/\mathbb{Q}_\ell\right)\right)$ contains g; in particular, for any such ℓ the group $\mathrm{Gal}\left(\overline{\mathbb{Q}_\ell}/\mathbb{Q}_\ell\right)$ acts transitively on the simple factors of $G_{\overline{\mathbb{Q}_\ell}}$, so $G_{\mathbb{Q}_\ell}$ is $\mathbb{Q}_\ell$-simple. □

Lemma 6.3.3. *Let K be a finitely generated subfield of $\mathbb{C}$ and A, B be K-abelian varieties. Suppose B is CM and $H(A \times B) \cong H(A) \times H(B)$. Then we have $H_\ell(A \times B) \cong H_\ell(A) \times H_\ell(B)$ for every prime ℓ.*

Proof. Using the hypothesis and applying theorem 6.2.14 twice we find

$$\begin{aligned}\mathrm{rk}\, Z\left(H_\ell\left(A \times B\right)\right) &= \mathrm{rk}\, Z\left(H\left(A \times B\right)\right) \\ &= \mathrm{rk}\, Z\left(H\left(A\right)\right) + \mathrm{rk}\, Z\left(H\left(B\right)\right) \\ &= \mathrm{rk}\, Z\left(H_\ell\left(A\right)\right) + \mathrm{rk}\, Z\left(H_\ell\left(B\right)\right).\end{aligned}$$

Furthermore, as $H_\ell(B)$ is a torus, the canonical projection $H_\ell(A\times B)\to H_\ell(A)$ induces an isogeny $H_\ell(A\times B)^{\mathrm{der}}\cong H_\ell(A)^{\mathrm{der}}$, hence $\operatorname{rk} H_\ell(A\times B)^{\mathrm{der}}=\operatorname{rk} H_\ell(A)^{\mathrm{der}}$. Putting these facts together we get $\operatorname{rk} H_\ell(A\times B)=\operatorname{rk} H_\ell(A)+\operatorname{rk} H_\ell(B)$, so the inclusion $H_\ell(A\times B)\hookrightarrow H_\ell(A)\times H_\ell(B)$ is an isomorphism by lemma 6.3.1. □

The next lemma is certainly well-known to experts (a somewhat similar statement is for example [116, Théorème 7], which deals with the case of elliptic curves), but for lack of an accessible reference we include a short proof:

Lemma 6.3.4. *Let K be a finitely generated subfield of $\mathbb{C}$ and A,B be K-abelian varieties. Suppose B is of CM type and $A_{\overline{K}}$ has no simple factor of type IV. Then we have $H(A\times B)\cong H(A)\times H(B)$, and for every prime ℓ we also have $H_\ell(A\times B)\cong H_\ell(A)\times H_\ell(B)$.*

Proof. The same proof works for both the Hodge group and the group H_ℓ, so let us only treat the former. The canonical projections $H(A\times B)\to H(A)$ and $H(A\times B)\to H(B)$ induce isogenies $H(A\times B)^{\mathrm{der}}\cong H(A)^{\mathrm{der}}$ and $Z(H(A\times B))\cong Z(H(B))$, so we have

$$\operatorname{rk} H(A\times B)=\operatorname{rk} H(A\times B)^{\mathrm{der}}+\operatorname{rk} Z(H(A\times B))=\operatorname{rk} H(A)^{\mathrm{der}}+\operatorname{rk} Z(H(B))=\operatorname{rk} H(A)+\operatorname{rk} H(B)$$

and we conclude by lemma 6.3.1. □

Lemma 6.3.5. *Let K be a finitely generated subfield of $\mathbb{C}$ and A,B be K-abelian varieties. Suppose that Mumford-Tate holds for A, and that B is CM. Then Mumford-Tate holds for $A\times B$.*

Proof. Let ℓ be a prime number. As in the previous lemma we have $\operatorname{rk} H_\ell(A\times B)^{\mathrm{der}}=\operatorname{rk} H_\ell(A)^{\mathrm{der}}$ and $\operatorname{rk} H(A\times B)^{\mathrm{der}}=\operatorname{rk} H(A)^{\mathrm{der}}$. Since the Mumford-Tate conjecture holds for A, we deduce $\operatorname{rk} H_\ell(A\times B)^{\mathrm{der}}=\operatorname{rk} H_\ell(A)^{\mathrm{der}}=\operatorname{rk} H(A)^{\mathrm{der}}=\operatorname{rk} H(A\times B)^{\mathrm{der}}$, and the lemma follows from proposition 6.2.16. □

Lemma 6.3.6. *Let K be a finitely generated subfield of $\mathbb{C}$ and $A_1,\dots,A_n$ be K-abelian varieties. Suppose that Mumford-Tate holds for every A_i, and that the equality $H_\ell\left(\prod_{i=1}^n A_i\right)=\prod_{i=1}^n H_\ell(A_i)$ holds for a given prime ℓ. Then the Mumford-Tate conjecture holds for $\prod_{i=1}^n A_i$.*

Proof. The hypothesis implies

$$\operatorname{rk} H_\ell\left(\prod_{i=1}^n A_i\right)=\sum_{i=1}^n \operatorname{rk} H_\ell(A_i)=\sum_{i=1}^n \operatorname{rk} H(A_i)\geq \operatorname{rk} H\left(\prod_{i=1}^n A_i\right),$$

and the lemma follows from proposition 6.2.16. □

One of the most important ingredients in our proofs is the following lemma, part of which is originally due to Ribet. The statement we give here is close in spirit to [80, Lemma 2.14], but our version is even more general.

Lemma 6.3.7. *Let $\mathbf{C}$ be an algebraically closed field of characteristic zero and $V_1,\dots,V_n$ be finite-dimensional $\mathbf{C}$-vector spaces. Let $\mathfrak{gl}(V_i)$ be the Lie algebra of endomorphisms of V_i and let $\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{gl}(V_1)\oplus\cdots\oplus\mathfrak{gl}(V_n)$. For each $i=1,\cdots,n$ let $\pi_i:\bigoplus_{j=1}^n\mathfrak{gl}(V_j)\to\mathfrak{gl}(V_i)$ be the i-th projection and let $\mathfrak{g}_i=\pi_i(\mathfrak{g})$. Suppose that each $\mathfrak{g}_i$ is a simple Lie algebra and that one of the following conditions holds:*

(a) For every pair of distinct indices i, j the projection $\pi_i \oplus \pi_j : \mathfrak{g} \to \mathfrak{g}_i \oplus \mathfrak{g}_j$ is onto.

(b) For all indices $i \neq j$ for which there is an isomorphism $\varphi : \mathfrak{g}_i \to \mathfrak{g}_j$ we have the following:

1. *there is an irreducible $\mathfrak{g}_i$-representation W such that all simple $\mathfrak{g}_i$-submodules of V_i and of $\varphi^*(V_j)$ are isomorphic to W, and the highest weight defining W is stable under all automorphisms of $\mathfrak{g}_i$;*
2. *let $I = \{k \in \{1, \dots, n\} \mid \mathfrak{g}_k \cong \mathfrak{g}_i\}$; the equality $\mathrm{End}_{\mathfrak{g}}\left(\bigoplus_{k\in I} V_k\right) \cong \prod_{k\in I} \mathrm{End}_{\mathfrak{g}_k} V_k$ holds.*

Then $\mathfrak{g} = \bigoplus_{j=1}^{n} \mathfrak{g}_j$.

Remark 6.3.8. As inner automorphisms preserve every highest weight, in condition (b1) one only needs to check the action of the outer automorphisms (which are finite in number, up to inner automorphisms, since they correspond to automorphisms of the Dynkin diagram). In particular, our conditions (b) generalize those given in [80, Lemma 2.14].

Proof. The fact that (a) implies the desired equality is classical, cf. the Lemma on pages 790-791 of [109]. Thus it suffices to show that (b) implies (a). Let us fix a pair (i, j) and consider the projection $\pi_i \oplus \pi_j : \mathfrak{g} \to \mathfrak{g}_i \oplus \mathfrak{g}_j$. Let $\mathfrak{h}$ be the image of this projection and $\mathfrak{k}$ be $\ker(\mathfrak{h} \to \mathfrak{g}_i)$. Since $\mathfrak{k}$ can be identified to an ideal of $\mathfrak{g}_j$ (which is simple), we either have $\mathfrak{k} \cong \mathfrak{g}_j$, in which case $\mathfrak{h} \cong \mathfrak{g}_i \oplus \mathfrak{g}_j$ as required, or $\mathfrak{k} = \{0\}$, in which case $\mathfrak{h}$ is the graph of an isomorphism $\mathfrak{g}_i \cong \mathfrak{g}_j$; it is this latter possibility that we need to exclude. If $\mathfrak{g}_i$ and $\mathfrak{g}_j$ are not isomorphic there is nothing to prove, so let us assume $\mathfrak{g}_i \cong \mathfrak{g}_j$, and suppose by contradiction that $\mathfrak{h}$ is the graph of an isomorphism $\varphi : \mathfrak{g}_i \to \mathfrak{g}_j$. Let $\rho_i : \mathfrak{g}_i \to \mathfrak{gl}(V_i)$ and $\rho_j : \mathfrak{g}_j \to \mathfrak{gl}(V_j)$ be the tautological representations of $\mathfrak{g}_i, \mathfrak{g}_j$. By assumption (b1), the simple $\mathfrak{g}_i$-subrepresentations of ρ_i and $\rho_j \circ \varphi$ are isomorphic, so there exists a nonzero morphism of $\mathfrak{g}_i$-representations $\chi_{ij} : V_i \to V_j$. Equivalently, χ_{ij} is $\mathfrak{h}$-equivariant (recall that $\mathfrak{h}$ is the graph of φ). Setting $I = \{k \in \{1, \dots, n\} \mid \mathfrak{g}_k \cong \mathfrak{g}_i\}$, the map

$$\begin{array}{rccc} \Psi : & \bigoplus_{k\in I} V_k & \to & \bigoplus_{k\in I} V_k \\ & (v_{i_1}, \cdots, \underbrace{v_i}_{\text{factor } V_i}, \cdots, v_{i_{|I|}}) & \mapsto & (0, \cdots, \underbrace{\chi_{ij}(v_i)}_{\text{factor } V_j}, \cdots, 0) \end{array}$$

then belongs to $\mathrm{End}_{\mathfrak{g}}\left(\bigoplus_{k\in I} V_k\right)$, but does not send every factor to itself, so it is not an element of $\prod_{k\in I} \mathrm{End}_{\mathfrak{g}_k}(V_k)$. This contradicts condition (b2), so $\mathfrak{g} \to \mathfrak{g}_i \oplus \mathfrak{g}_j$ must be onto, and therefore (b) implies (a) as required. □

Proposition 6.3.9. *Let K be a finitely generated field of characteristic zero, A, B be K-abelian varieties and ℓ be a prime number. Suppose at least one among $H_\ell(A)$ and $H_\ell(B)$ is semisimple, and no simple factor of $\mathrm{Lie}(H_\ell(A))^{\mathrm{ss}} \otimes \overline{\mathbb{Q}_\ell}$ is isomorphic to a simple factor of $\mathrm{Lie}(H_\ell(B))^{\mathrm{ss}} \otimes \overline{\mathbb{Q}_\ell}$: then $\mathrm{H}_\ell(A \times B) \cong H_\ell(A) \times H_\ell(B)$.*

Proof. Up to interchanging A and B we can assume that $H_\ell(A)$ is semisimple: the projection $H_\ell(A \times B) \twoheadrightarrow H_\ell(B)$ then induces an isogeny $Z(H_\ell(A \times B)) \cong Z(H_\ell(B))$.

Next consider the semisimple ranks. Let $\mathfrak{h}$, $\mathfrak{h}_A$ and $\mathfrak{h}_B$ be the Lie algebras $\mathrm{Lie}(H_\ell(A\times B))^{\mathrm{ss}}\otimes\overline{\mathbb{Q}_\ell}$, $\mathrm{Lie}(H_\ell(A))\otimes\overline{\mathbb{Q}_\ell}$ and $\mathrm{Lie}(H_\ell(B))^{\mathrm{ss}}\otimes\overline{\mathbb{Q}_\ell}$ respectively.
Write $\mathfrak{h}_A\cong\mathfrak{g}_1\oplus\cdots\oplus\mathfrak{g}_n$ and $\mathfrak{h}_B\cong\mathfrak{g}_{n+1}\oplus\cdots\oplus\mathfrak{g}_{n+m}$, with every $\mathfrak{g}_i$ simple. We can consider $\mathfrak{h}$ as a subalgebra of $\bigoplus_{i=1}^n\mathfrak{g}_i\oplus\bigoplus_{j=1}^m\mathfrak{g}_{n+j}$ that projects surjectively onto $\bigoplus_{i=1}^n\mathfrak{g}_i$ and $\bigoplus_{j=1}^m\mathfrak{g}_{n+j}$. In particular, $\mathfrak{h}$ projects surjectively onto each simple factor $\mathfrak{g}_i$.
Let us show that all the double projections $\mathfrak{h}\to\mathfrak{g}_i\oplus\mathfrak{g}_j$ are onto. If i,j are both at most n (or i,j are both at least $n+1$) this is trivial, so we can assume $i\le n<j$. But then by assumption $\mathfrak{g}_i$ and $\mathfrak{g}_j$ are nonisomorphic, so by the same argument as in the proof of lemma 6.3.7 the projection must be surjective. Lemma 6.3.7 now gives $\mathfrak{h}\cong\mathfrak{h}_A\oplus\mathfrak{h}_B$, thus implying $\operatorname{rk}\mathfrak{h}=\operatorname{rk}\mathfrak{h}_A+\operatorname{rk}\mathfrak{h}_B$. In terms of groups this leads to

$$\begin{aligned}\operatorname{rk}H_\ell(A\times B)&=\operatorname{rk}H_\ell(A\times B)^{\mathrm{der}}+\operatorname{rk}Z(H_\ell(A\times B))\\&=\operatorname{rk}H_\ell(A)^{\mathrm{der}}+\operatorname{rk}H_\ell(B)^{\mathrm{der}}+\operatorname{rk}Z(H_\ell(B))\\&=\operatorname{rk}H_\ell(A)+\operatorname{rk}H_\ell(B),\end{aligned}$$

and we conclude by lemma 6.3.1. □

6.4 Sufficient conditions for H_ℓ to decompose as a product

6.4.1 An ℓ-adic analogue of a theorem of Hazama

We are now ready to prove the following ℓ-adic analogue (and mild generalization) of a Hodge-theoretical result of Hazama ([36, Proposition 1.8]):

Theorem 6.4.1. *Let K be a finitely generated field of characteristic zero, A_1 and A_2 be K-abelian varieties, and ℓ be a prime number. For $i=1,2$ let $\mathfrak{h}_i$ be the Lie algebra of $H_\ell(A_i)$. Suppose that the following hold:*

1. *for $i=1,2$, the algebra $\mathfrak{h}_i$ is semisimple, so that we can write $\mathfrak{h}_i\otimes\overline{\mathbb{Q}_\ell}\cong\mathfrak{h}_{i,1}\oplus\cdots\oplus\mathfrak{h}_{i,n_i}$, where every $\mathfrak{h}_{i,j}$ is simple;*
2. *for $i=1,2$, there exists a decomposition $V_\ell(A_i)\otimes\overline{\mathbb{Q}_\ell}\cong V_{i,1}\oplus\cdots\oplus V_{i,n_i}$ such that the action of $\mathfrak{h}_i\otimes\overline{\mathbb{Q}_\ell}\cong\mathfrak{h}_{i,1}\oplus\cdots\oplus\mathfrak{h}_{i,n_i}$ on $V_{i,1}\oplus\cdots\oplus V_{i,n_i}$ is componentwise and $\mathfrak{h}_{i,j}$ acts faithfully on $V_{i,j}$;*
3. *for all distinct pairs (i,j) and (i',j') for which there exists an isomorphism $\varphi:\mathfrak{h}_{i,j}\to\mathfrak{h}_{i',j'}$ there is an irreducible $\mathfrak{h}_{i,j}$-representation W such that all simple $\mathfrak{h}_{i,j}$-submodules of $V_{i,j}$ and of $\varphi^*\left(V_{i',j'}\right)$ are isomorphic to W, and the highest weight defining W is stable under all automorphisms of $\mathfrak{h}_{i,j}$.*

Then either $\mathrm{Hom}_{\overline{K}}(A_1,A_2)\neq 0$ *or* $H_\ell(A_1\times A_2)\cong H_\ell(A_1)\times H_\ell(A_2)$.

Remark 6.4.2. Condition 3, as W is fixed under all automorphisms of $\mathfrak{g}_{i,j}$ (hence of $\mathfrak{g}_{i',j'}$), is actually independent of the choice of φ.

Proof. Let $\mathfrak{h}$ be the Lie algebra of $H_\ell(A_1 \times A_2)$. We shall try to apply lemma 6.3.7 to the inclusion $\mathfrak{h} \otimes \overline{\mathbb{Q}_\ell} \hookrightarrow (\mathfrak{h}_1 \oplus \mathfrak{h}_2) \otimes \overline{\mathbb{Q}_\ell}$, and distinguish cases according to whether hypothesis (b2) is satisfied or not. Observe that $\mathfrak{h} \otimes \overline{\mathbb{Q}_\ell}$ is a subalgebra of

$$(\mathfrak{h}_1 \oplus \mathfrak{h}_2) \otimes \overline{\mathbb{Q}_\ell} \cong \bigoplus_{i=1}^{2} \bigoplus_{j=1}^{n_i} \mathfrak{h}_{i,j} \subset \bigoplus_{i=1}^{2} \bigoplus_{j=1}^{n_i} \mathfrak{gl}\left(V_{i,j}\right)$$

whose projection on each factor $\mathfrak{gl}\left(V_{i,j}\right)$ is isomorphic to $\mathfrak{h}_{i,j}$, hence simple. Moreover, hypothesis 3 of this theorem implies condition (b1) of lemma 6.3.7. Suppose now that (b2) holds as well: then $\mathfrak{h} \otimes \overline{\mathbb{Q}_\ell} \cong (\mathfrak{h}_1 \oplus \mathfrak{h}_2) \otimes \overline{\mathbb{Q}_\ell}$, hence in particular $\operatorname{rk} \mathfrak{h} = \operatorname{rk} \mathfrak{h}_1 + \operatorname{rk} \mathfrak{h}_2$, and lemma 6.3.1 implies $H_\ell(A_1 \times A_2) \cong H_\ell(A_1) \times H_\ell(A_2)$. Suppose on the other hand that (b2) fails: then there exists a nontrivial endomorphism φ in

$$\operatorname{End}_{\mathfrak{h} \otimes \overline{\mathbb{Q}_\ell}} \left(\bigoplus_{i=1}^{2} \bigoplus_{j=1}^{n_i} V_{i,j} \right) \setminus \bigoplus_{i=1}^{2} \bigoplus_{j=1}^{n_i} \operatorname{End}_{\mathfrak{h}_{i,j}} \left(V_{i,j}\right).$$

Since the action of $\mathfrak{h}_i \otimes \overline{\mathbb{Q}_\ell}$ on $V_\ell(A_i) \otimes \overline{\mathbb{Q}_\ell} \cong \bigoplus_{j=1}^{n_i} V_{i,j}$ is componentwise for $i = 1, 2$, it is clear that φ does not belong to $\operatorname{End}_{\mathfrak{h}_1}\left(\bigoplus_{j=1}^{n_1} V_{1,j}\right) \times \{0\}$, nor to $\{0\} \times \operatorname{End}_{\mathfrak{h}_2}\left(\bigoplus_{j=1}^{n_2} V_{2,j}\right)$. Thus, up to exchanging the roles of A_1 and A_2 if necessary, the map φ induces an $(\mathfrak{h} \otimes \overline{\mathbb{Q}_\ell})$-equivariant morphism from $\bigoplus_{j=1}^{n_1} V_{1,j}$ to $\bigoplus_{j=1}^{n_2} V_{2,j}$: this implies that the space

$$\operatorname{Hom}_{\mathfrak{h}}\left(V_{\ell,1}, V_{\ell,2}\right) \otimes \overline{\mathbb{Q}_\ell} \cong \operatorname{Hom}_{\mathfrak{h} \otimes \overline{\mathbb{Q}_\ell}}\left(V_{\ell,1} \otimes \overline{\mathbb{Q}_\ell}, V_{\ell,2} \otimes \overline{\mathbb{Q}_\ell}\right)$$

is nontrivial. In particular, $\operatorname{Hom}_{\mathfrak{h}}\left(V_\ell(A_1), V_\ell(A_2)\right) \neq 0$, and therefore $\operatorname{Hom}_{\overline{K}}(A_1, A_2)$ is nontrivial by corollary 6.2.7. □

Remark 6.4.3. We now check to what extent the theorem can be applied to varieties A of general Lefschetz type. It is clear that conditions 1 and 2 are satisfied, so let us discuss condition 3. Let $\mathfrak{h}$ be a simple constituent of $\operatorname{Lie} H_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$. By definition, the simple $\mathfrak{h}$-submodules of $V_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ are all isomorphic to a single representation W. Let us distinguish cases according to the type of $\mathfrak{h}$:

- if $\mathfrak{h}$ is of Lie type A_l, then W is defined by the highest weight $\omega_{\frac{l+1}{2}}$ (recall that l is odd by assumption), and is therefore stable under the unique nontrivial automorphism of the Dynkin diagram of A_l: condition 3 is satisfied;
- if $\mathfrak{h}$ is of Lie type B_l or C_l, the Dynkin diagram does not have any nontrivial automorphisms, hence all automorphisms of $\mathfrak{h}$ are inner and fix the highest weight of W: condition 3 is again satisfied;
- finally, if $\mathfrak{h}$ is of Lie type D_l the module W is defined by the highest weight ω_1. As long as $l \neq 4$, the Dynkin diagram of D_l has a unique nontrivial automorphism, and it is immediate to check that this automorphism fixes ω_1: condition 3 is satisfied once more. Note however that for $l = 4$ the Dynkin diagram has additional (triality) automorphisms, and that these do *not* fix ω_1, so condition 3 fails in this case.

Thus we conclude that every abelian variety A of general Lefschetz type satisfies the hypotheses of the previous theorem unless $\operatorname{Lie} H_\ell(A) \otimes \overline{\mathbb{Q}_\ell}$ has a simple factor of Lie type D_4.

Corollary 6.4.4. *Let K be a finitely generated subfield of $\mathbb{C}$ and $A_1, \dots, A_n$ be absolutely simple abelian varieties defined over K, pairwise non-isogenous over $\overline{K}$. Suppose that no A_i is of type IV, and that the dimension of each A_i is either 2 or an odd number. Let $k_1, \dots, k_n$ be positive integers and A be a K-abelian variety that is $\overline{K}$-isogenous to $\prod_{i=1}^n A_i^{k_i}$. Then we have an isomorphism $H_\ell(A) \cong \prod_{i=1}^n H_\ell(A_i)$, and the Mumford-Tate conjecture holds for A.*

Proof. The Albert classification implies that every A_i is of type I or II (recall that in characteristic zero there is no absolutely simple abelian surface of type III). As the three abelian varieties $\prod_{i=1}^n A_i^{k_i}$, $\prod_{i=1}^n A_i$ and A all have the same Hodge group and the same groups H_ℓ, there is no loss of generality in assuming that $k_1 = \cdots = k_n = 1$ and that $A = \prod_{i=1}^n A_i$. The fact that $H_\ell(A_1 \times \cdots \times A_n)$ and $H_\ell(A_1) \times \cdots \times H_\ell(A_n)$ are isomorphic then follows by induction from theorem 6.4.1, the hypotheses being satisfied thanks to theorem 6.2.24 (and the remark following it). Lemma 6.3.6 then implies that Mumford-Tate holds for $A_1 \times \cdots \times A_n$. □

Corollary 6.4.5. *Let K be a finitely generated subfield of $\mathbb{C}$ and $A_1, \dots, A_n$ be absolutely simple K-abelian varieties of dimension at most 2, pairwise non-isogenous over $\overline{K}$. Let $k_1, \dots, k_n$ be positive integers and A be a K-abelian variety that is $\overline{K}$-isogenous to $\prod_{i=1}^n A_i^{k_i}$. Then we have $H_\ell(A) \cong \prod_{i=1}^n H_\ell(A_i)$, and the Mumford-Tate conjecture holds for A.*

Remark 6.4.6. Such a result is in a sense the best possible. There is an example – due to Shioda [127] – of an absolutely simple threefold Y of CM type and a CM elliptic curve E such that $H(Y \times E) \neq H(Y) \times H(E)$. By the Mumford-Tate conjecture in the CM case, this also means $H_\ell(Y \times E) \neq H_\ell(Y) \times H_\ell(E)$ (note that Y and E, being CM, can be defined over a number field).

Proof. As in the previous proof, we can assume $k_1 = \cdots = k_n = 1$ and replace A by $\prod_{i=1}^n A_i$. By lemma 6.3.6, Mumford-Tate for A would follow from the isomorphism $H_\ell(A) \cong \prod_{i=1}^n H_\ell(A_i)$, so let us prove the latter. Up to renumbering, we can also assume that $A_1, \dots, A_m$ are of type I or II and $A_{m+1}, \dots, A_n$ are of type IV (since there are no absolutely simple abelian varieties of type III of dimension at most 2). The classification of elliptic curves and simple surfaces implies that $A_{m+1}, \dots, A_n$ are CM. Let $A' = A_1 \times \cdots \times A_m$ and $A'' = A_{m+1} \times \cdots \times A_n$. As A'' is CM and A' has no simple factor of type IV, lemma 6.3.4 gives $H_\ell(A' \times A'') \cong H_\ell(A') \times H_\ell(A'')$. It thus suffices to prove the result when either A' or A'' is trivial. If A'' is trivial the claim follows from corollary 6.4.4, so we can assume A' is trivial, in which case we have to show $H_\ell(\prod_{i=1}^n A_i) \cong \prod_{i=1}^n H_\ell(A_i)$ under the additional assumption that every A_i is CM. Appealing to the Mumford-Tate conjecture in the CM case, it is enough to show the corresponding statement for Hodge groups, which is exactly the content of [103, Theorem 3.15]. □

6.4.2 A criterion in terms of relative dimensions

As promised in the introduction, we have the following ℓ-adic analogue of a theorem proved by Ichikawa in [44]:

Theorem 6.4.7. *Let K be a finitely generated field of characteristic zero and A'_i, A''_j (for $i = 1, \dots, n$ and $j = 1, \dots, m$) be absolutely simple K-abelian varieties of odd relative dimension that are pairwise*

non-isogenous over $\overline{K}$*. Suppose every* A'_i *is of type I, II or III in the sense of Albert, and every* A''_j *is of type IV. Let* A *be a* K*-abelian variety that is* $\overline{K}$*-isogenous to* $\prod_{i=1}^n A'_i \times \prod_{j=1}^m A''_j$*: then*

$$H_\ell(A) \cong \prod_{i=1}^{n} H_\ell\left(A'_i\right) \times H_\ell\left(\prod_{j=1}^{m} A''_j\right).$$

For the proof of this theorem we shall need the following result:

Proposition 6.4.8. *Let* K *be a finitely generated field of characteristic zero,* A/K *be an absolutely simple abelian variety of odd relative dimension and* ℓ *be a prime number. Write* $\mathrm{Lie}(H_\ell(A)) \otimes \overline{\mathbb{Q}_\ell}$ *as* $\mathfrak{c} \oplus \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$*, where* $\mathfrak{c}$ *is abelian and every* $\mathfrak{h}_i$ *is simple. Then*

1. *if* A *is of type I, II or III, then* A *satisfies all the hypotheses of theorem 6.4.1;*
2. *if* A *is of type IV, then the algebras* $\mathfrak{h}_i$ *are of type* A_l*, where* $l+1$ *is not a power of* 2*.*

Proof. Let A be of type I, II or III. Then A is of general Lefschetz type by theorem 6.2.24 and proposition 6.2.26, and again by proposition 6.2.26 the simple factors of $\mathrm{Lie}\,(H_\ell(A)) \otimes \overline{\mathbb{Q}_\ell}$ of orthogonal type are of the form $\mathfrak{so}_{2h}$ with h odd, so none of them is of Lie type D_4. Hence A satisfies the hypotheses of theorem 6.4.1 by remark 6.4.3.
Let now A be of type IV. Let E be the center of the simple algebra $\mathrm{End}^0_{\overline{K}}(A)$; set $e = [E:\mathbb{Q}]$ and $d^2 = \left[\mathrm{End}^0_{\overline{K}}(A) : E\right]$. We are first going to show the desired property for those primes that split in E, and then extend the result to all primes through an interpolation argument based on the techniques of [57]. Suppose therefore that ℓ is totally split in E. From the equality $E \otimes \mathbb{Q}_\ell \cong \mathbb{Q}_\ell^{[E:\mathbb{Q}]}$ we get

$$\mathrm{End}^0_{\overline{K}}(A) \otimes \overline{\mathbb{Q}_\ell} \cong \bigoplus_{\sigma: E \hookrightarrow \mathbb{C}} M_d(\overline{\mathbb{Q}_\ell}),$$

so Schur's lemma implies

$$V_\ell(A) \otimes \overline{\mathbb{Q}_\ell} \cong \bigoplus_{\sigma: E \hookrightarrow \mathbb{C}} W_\sigma^{\oplus d},$$

where each W_σ is simple of dimension $\frac{1}{de}\dim_{\overline{\mathbb{Q}_\ell}}(V_\ell(A) \otimes \overline{\mathbb{Q}_\ell}) = \mathrm{reldim}(A)$. The action of $H_\ell(A)$ on $V_\ell(A)$ is faithful, so for every $i = 1, \ldots, n$ there exists a $\sigma : E \hookrightarrow \mathbb{C}$ (depending on i) such that the action of $\mathfrak{h}_i$ is nontrivial on W_σ. Note that $\dim(W_\sigma)$ is odd. Let $W_\sigma \cong Z_1 \otimes \cdots \otimes Z_n$ be the decomposition of W_σ with respect to the action of $\mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$; the module Z_i is thus a nontrivial minuscule representation of $\mathfrak{h}_i$ of *odd* dimension: since every minuscule module over an algebra of type B_l, C_l, D_l is of *even* dimension (cf. table 1), we deduce that $\mathfrak{h}_i$ is of type A_l for a certain l. Furthermore, $l+1$ cannot be a power of 2, since in that case every irreducible minuscule module over A_l is of even dimension. This shows our claim when ℓ is totally split.

Let us now consider the general case. Let ℓ be any prime, and p be a fixed prime that splits completely in E. Let Φ_ℓ be the root system of $\left(G_\ell(A) \otimes \overline{\mathbb{Q}_\ell}\right)^{\mathrm{der}}$, and let Φ_ℓ^0 be the subset of Φ_ℓ given by those roots that are short in their respective simple factors of $\left(G_\ell(A) \otimes \overline{\mathbb{Q}_\ell}\right)^{\mathrm{der}}$. Note that $\Phi_p^0 = \Phi_p$, since Φ_p only involves root systems of type A_l (and such root systems do not possess long roots). It is a theorem of Serre that the formal characters of the various $G_\ell(A)$, for varying ℓ, are all equal (see [98, Corollary 3.8]), and from [57, §4] (see also pp. 212-213 of [98]) we know

that the formal character completely determines Φ_ℓ^0. Hence we have $\Phi_\ell^0 = \Phi_p^0 = \bigoplus_{i=1}^k A_{n_i}$ for a certain k and for integers n_i such that no $n_i + 1$ is a power of 2; in particular, no n_i equals 1. Write now $\Phi_\ell = \bigoplus_{i=1}^r R_i$, where each R_i is a simple root system. It is easy to see that $A_l^0 = A_l, B_l^0 = lA_1, C_l^0 = D_l$ and $D_l^0 = D_l$, so the equality

$$\bigoplus_{i=1}^{k} A_{n_i} = \Phi_p^0 = \Phi_\ell^0 = \bigoplus_{j=1}^{r} R_j^0$$

implies – by uniqueness of the decomposition in simple root systems – that every root system R_j is either of type A_l or B_m (for some l, m). On the other hand, if one R_j were of type B_m, then the right hand side of the above equality would contain $B_m^0 = mA_1$, but no root system of type A_1 can appear on the left hand side by what we have already shown. This implies that every R_j is of type A_l (for some l), and the uniqueness of the decomposition shows that $r = k$ and (up to renumbering the indices) $R_j = A_{n_j}$. Hence the root system of $G_\ell(A)^{\mathrm{der}}$ is the same as that of $G_p(A)^{\mathrm{der}}$, and in particular all the simple algebras $\mathfrak{h}_i$ are of Lie type A_l, where $l + 1$ is not a power of 2. □

Proof. (of Theorem 6.4.7) There is no loss of generality in assuming that $A = A' \times A''$, where

$$A' = \prod_{i=1}^{n} A_i', \quad A'' = \prod_{j=1}^{m} A_j''.$$

Thanks to the previous proposition, theorem 6.4.1 and an immediate induction imply that $H_\ell(A')$ is isomorphic to $\prod_{i=1}^n H_\ell(A_i')$. Thus it is enough to show that $H_\ell(A) \cong H_\ell(A') \times H_\ell(A'')$, and this follows from proposition 6.3.9: by the results of section 6.2.4, the simple factors of $\mathrm{Lie}\,(H_\ell(A')) \otimes \overline{\mathbb{Q}_\ell}$ are either of type $\mathfrak{so}$, $\mathfrak{sp}$ or $\mathfrak{sl}_{l+1}$ (with $l + 1$ a power of 2), whereas by the previous proposition the simple factors of $\mathrm{Lie}\left(H_\ell(A'')^{\mathrm{der}}\right) \otimes \overline{\mathbb{Q}_\ell}$ are of type $\mathfrak{sl}_{l+1}$ (with $l + 1$ not a power of 2). □

Remark 6.4.9. Notice that, as the rank of $H_\ell(A)$ is independent of ℓ, knowing that part (2) of proposition 6.4.8 holds for *some* prime ℓ would in fact be enough to prove theorem 6.4.7. Though a weaker version of the proposition would be easier to show (since it would not require the second part of the proof provided), we have preferred to give and employ the result in its stronger form (applying to *all* primes), which we believe has some merit in itself.

6.5 Results in positive characteristic

We now discuss the situation of K being a field of positive characteristic, finitely generated over its prime field, and we restrict ourselves to the primes $\ell \neq \mathrm{char}\, K$. If A is a K-abelian variety, we denote $G_\ell(A)$ the Zariski closure of the natural Galois representation

$$\rho_\ell : \mathrm{Gal}\,(K^{\mathrm{s}}/K) \to \mathrm{Aut}\,(T_\ell(A))\,,$$

where K^{s} is now a fixed *separable* closure of K.

The main difficulty in translating the results of the previous sections to this context is that if we define $H_\ell(A)$ as $(G_\ell(A) \cap \mathrm{SL}(V_\ell(A)))^0$, then this group might not capture any information about A at all. The crucial problem is the failure of Bogomolov's theorem in positive characteristic: for general abelian varieties A/K, it is not true that $G_\ell(A)$ contains the torus of homotheties, and therefore the intersection $G_\ell(A) \cap \mathrm{SL}(V_\ell(A))$ may very well be finite.

Remark 6.5.1. A simple example of this phenomenon is given by an ordinary elliptic curve E over a finite field $\mathbb{F}_q$. Let Fr_q be the Frobenius automorphism of $\mathbb{F}_q$; the image of ρ_ℓ is generated by the image g of Fr_q, and as it is well known we have $\det \rho_\ell(g) = q$. Looking at the Lie algebra of $G_\ell(E)$, it follows easily that this group is 1-dimensional and that $H_\ell(E)$ is the trivial group, so that no information about E can be recovered from $H_\ell(E)$. This problem is studied in [154], where more examples of this situation are given.

However, Zarhin has proved that a statement akin to Bogomolov's theorem holds in positive characteristic if we restrict ourselves to a certain (large) class of abelian varieties; more precisely, we have the following result:

Theorem 6.5.2. *([150], Theorem 2 and Corollary 1) Let K be a finitely generated field of positive characteristic and A be a K-abelian variety. Let ℓ be a prime different from $\mathrm{char}(K)$. There exist a semisimple Lie algebra $\mathfrak{h}$ and a 1-dimensional Lie algebra $\mathfrak{c}$ such that $\mathrm{Lie}\, G_\ell(A) \cong \mathfrak{c} \oplus \mathfrak{h}$.*
If furthermore no simple factor of $A_{\overline{K}}$ is of type IV in the sense of Albert, then $\mathfrak{c} \cong \mathbb{Q}_\ell \cdot \mathrm{Id}$ is the Lie algebra of the torus of homotheties.

Remark 6.5.3. Zarhin's theorem is a rather direct consequence of the reductivity of $G_\ell(A)$ and of Tate's conjecture on homomorphisms. At the time of [150], these two facts had only been established (by Zarhin himself, cf. [148] and [149]) under the assumption that $\mathrm{char}\, K$ is greater than 2, but Mori [83] has subsequently lifted this restriction.

Remark 6.5.4. Let K be a finitely generated field of positive characteristic and E_1, E_2 be two elliptic curves over K. Assume $\mathrm{End}_{\overline{K}}(E_1)$ and $\mathrm{End}_{\overline{K}}(E_1)$ are imaginary quadratic fields, and E_1, E_2 are not isogenous over $\overline{K}$. As $E_1 \times E_2$ is CM, the group $G_\ell(E_1 \times E_2)$ is abelian and therefore – by Zarhin's theorem – of dimension 1: this is in stark contrast with what happens in characteristic zero, where $H_\ell(E_1 \times E_2) \cong H_\ell(E_1) \times H_\ell(E_2)$ is of dimension 2. In particular, we cannot hope for an analogue of corollary 6.4.5 to hold in positive characteristic.

In view of Zarhin's theorem and of the previous remarks, the most natural definition for $H_\ell(A)$ in positive characteristic seems to be the following:

Definition 6.5.5. Let K be a finitely generated field of characteristic $p > 0$. For every prime ℓ different from p we set $H_\ell(A) = \left(G_\ell(A)^0\right)^{\mathrm{der}}$.

Remark 6.5.6. When the characteristic of K is positive, Zarhin's theorem implies that $G_\ell(A)^{\mathrm{der}}$ is of codimension 1 in $G_\ell(A)$; this is not necessarily the case in characteristic zero. On the other hand, as in characteristic zero, it is clear from definition 6.5.5 that $H_\ell(A \times B)$ projects surjectively onto $H_\ell(A)$ and $H_\ell(B)$.

Let us now restrict ourselves to abelian varieties A such that no simple factor of $A_{\overline{K}}$ is of type IV. In the proof of corollary 6.2.7 we can then replace Bogomolov's theorem by Zarhin's theorem, at which point the argument used to show theorem 6.4.1 goes through essentially unchanged. Thus for this class of abelian varieties we have:

Theorem 6.5.7. *(cf. theorem 6.4.1) Let K be a finitely generated field of characteristic $p > 0$ and A_1, A_2 be K-abelian varieties such that $A_{1,\overline{K}}$ and $A_{2,\overline{K}}$ have no simple factors of type IV. Let ℓ be a prime number different from p, and suppose hypotheses 1 through 3 of theorem 6.4.1 are satisfied. Then either $\mathrm{Hom}_{\overline{K}}(A_1, A_2) \neq 0$ or $H_\ell(A_1 \times A_2) \cong H_\ell(A_1) \times H_\ell(A_2)$.*

Remark 6.5.8. This theorem is strictly weaker than the corresponding result in characteristic zero, in that there exist abelian varieties of type IV (over number fields) that satisfy all hypotheses of theorem 6.4.1. Examples of such varieties include fourfolds of type IV(1,1) that support exceptional Weil classes, cf. [80]. On the other hand, the abelian varieties of corollary 6.4.4 satisfy the hypotheses of the present weakened version, hence the corollary remains true when K is of positive characteristic.

Let us now consider theorem 6.4.7. Its proof essentially relies on theorem 6.2.24 and proposition 6.2.26, which in turn only depend on Tate's conjecture and on the minuscule weights conjecture (theorem 6.2.9). As already remarked, the former is now known for arbitrary finitely generated fields of positive characteristic, while the second has been shown by Zarhin ([152, Theorem 4.2]) under an additional technical assumption, namely that the abelian variety in question has ordinary reduction in dimension 1 at all places of K with at most finitely many exceptions (cf. [152, Definition 4.1.0]; this is a condition weaker than being ordinary). Finally, for varieties of type IV we have also exploited the fact that the formal character of $G_\ell(A)^0$ is independent of ℓ: this statement too is known for finitely generated fields of positive characteristic (see [151] and [58], Proposition 6.12 and Examples 6.2, 6.3), so proposition 6.4.8 is still valid in this context. Taking all these facts into account we obtain:

Theorem 6.5.9. *(cf. theorem 6.4.7) Let K be a finitely generated field of positive characteristic and A'_i, A''_j (for $i = 1, \dots, n$ and $j = 1, \dots, m$) be absolutely simple K-abelian varieties of odd relative dimension that are pairwise non-isogenous over $\overline{K}$. Suppose every A'_i is of type I, II or III in the sense of Albert, and every A''_j is of type IV. Finally, suppose that each A'_i and each A''_j has ordinary reduction in dimension 1 at all places of K with at most finitely many exceptions, and let ℓ be a prime different from* $\operatorname{char} K$. *Let A be a K-abelian variety that is $\overline{K}$-isogenous to $\prod_{i=1}^n A'_i \times \prod_{j=1}^m A''_j$: then*

$$H_\ell(A) \cong \prod_{i=1}^{n} H_\ell\left(A'_i\right) \times H_\ell\left(\prod_{j=1}^{m} A''_j\right).$$

6.6 Nonsimple varieties of dimension at most 5

Let once more K be a finitely generated subfield of $\mathbb{C}$ and A/K be an abelian variety. With the results of the previous sections at hand it is a simple matter to compute, when A/K is of dimension at most 5 and nonsimple over $\overline{K}$, the structure of $H_\ell(A)$ in terms of the H_ℓ's of the simple factors of $A_{\overline{K}}$. Given however that the analogous problem for $H(A)$ has been given a complete solution in [82], we limit ourselves to showing that (in most cases) such an A satisfies Mumford-Tate, and refer the reader to [82] for more details on the precise structure of $H(A)$ (hence of $H_\ell(A)$). Note in any case that – for many varieties, including those for which we cannot prove Mumford-Tate – our argument will yield the structure of $H_\ell(A)$ directly, without appealing to the results of [82].

Proposition 6.6.1. *Let K be a finitely generated subfield of $\mathbb{C}$, n be an integer no less than 2, and $A_1, \dots, A_n$ be absolutely simple K-abelian varieties such that $\sum_{i=1}^n \dim A_i \leq 4$. Let A be a K-abelian variety that is $\overline{K}$-isogenous to $A_1 \times \cdots \times A_n$: then the Mumford-Tate conjecture holds for A.*

Proof. Since $H(A)$ and $H_\ell(A)$ are invariant both under isogeny and finite extension of the base field, we can assume without loss of generality that $A = A_1 \times \cdots \times A_n$. If all the A_i's are of dimension at most 2 we can simply apply corollary 6.4.5, so we can also assume that A_1 is an absolutely simple threefold and A_2 is an elliptic curve. In particular, A_1 and A_2 are of odd relative dimension, so if A_2 does not have complex multiplication (hence it is not of type IV) we have $H_\ell(A_1 \times A_2) \cong H_\ell(A_1) \times H_\ell(A_2)$ by theorem 6.4.7, and the claim follows from lemma 6.3.6. On the other hand, if A_2 does have complex multiplication the claim follows immediately from lemma 6.3.5. □

Proposition 6.6.2. *Let K be a finitely generated subfield of $\mathbb{C}$ and $A_1, \ldots, A_n$ be absolutely simple K-abelian varieties. Let A be a K-abelian variety that is $\overline{K}$-isogenous to $A_1 \times \cdots \times A_n$, and suppose that $\dim A = 5$. We have:*

1. *if Mumford-Tate holds for every A_i, then it also holds for A (this happens in particular if no A_i is of dimension 4);*
2. *if Mumford-Tate fails for one of the A_i's, say A_1, then A_1 is an absolutely simple fourfold, A_2 is an elliptic curve, and $H_\ell(A) = H_\ell(A_1 \times A_2) \cong H_\ell(A_1) \times H_\ell(A_2)$.*

Proof. We can work with $A = A_1 \times \cdots \times A_n$, and we can assume that no two A_i's are isogenous over $\overline{K}$ (for otherwise the problem is reduced to a lower-dimensional one). Furthermore, for $n = 1$ there is nothing to prove, so let us assume $n \geq 2$; as in the proof of the previous proposition, up to renumbering the A_i's we can assume that $\dim A_1 \geq 3$.
Suppose first that at least one of the A_i's has complex multiplication. Write $A = B \times C$, where C is the product of those A_i's that are CM and B is the product of the remaining factors. We have $\dim B \leq 4$. If B satisfies Mumford-Tate, then Mumford-Tate for A follows from lemma 6.3.5 and we are done. If, on the contrary, B does not satisfy Mumford-Tate, then the results of section 6.2.4 together with the previous proposition imply that $B = A_1$ is an absolutely simple fourfold with $\mathrm{End}_{\overline{K}}(B) = \mathbb{Z}$, and we are in case (2); hence we just need to prove that $H_\ell(A_1 \times A_2)$ is isomorphic to $H_\ell(A_1) \times H_\ell(A_2)$, which follows at once from lemma 6.3.4. From now on we can therefore assume that no A_i is CM. Also recall that elliptic curves and abelian surfaces without CM are of type I or II in the sense of Albert.
We now need to distinguish several sub-cases, each of which we shall treat by proving the equality $H_\ell(A) \cong \prod_{i=1}^n H_\ell(A_i)$: indeed, if Mumford-Tate holds for every A_i, this equality implies Mumford-Tate for A by lemma 6.3.6, and if Mumford-Tate fails for one of the A_i's this equality is all we have to show.

Suppose first that $\dim A_1 = 3$ and A_2, A_3 are elliptic curves (without CM): then for all primes ℓ, and independently of the type of A_1, theorem 6.4.7 gives $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2) \times H_\ell(A_3)$.

Next suppose $\dim A_1$ is 3 and A_2 is an absolutely simple abelian surface without CM (hence not of type IV). Let ℓ be any prime. If $\mathrm{reldim}(A_2) = 1$, or A_1 is not of type IV, then we have $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2)$ resp. by theorem 6.4.7 or corollary 6.4.4. We can therefore assume that $\mathrm{End}_{\overline{K}}(A_2)$ is $\mathbb{Z}$ and A_1 is of type IV and does not have complex multiplication. It is known that in this case $\mathrm{Lie}(H_\ell(A_2)) \cong \mathfrak{sp}_{4,\mathbb{Q}_\ell}$, and $\mathrm{Lie}\left(H_\ell(A_1)^{\mathrm{der}}\right) \otimes \overline{\mathbb{Q}_\ell} \cong \mathfrak{sl}_{3,\overline{\mathbb{Q}_\ell}}$ (cf. [112]), so it follows from proposition 6.3.9 that $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2)$.

We now need to consider the case when A_1 is an absolutely simple abelian fourfold and A_2 is an elliptic curve without CM; this assumption will be in force for the remainder of the proof.
Suppose first that A_1 is not of type IV and that $\mathrm{End}_{\overline{K}}(A_1) \neq \mathbb{Z}$. By the results of [80] we know that A_1 is of general Lefschetz type, so that the equality $H_\ell(A_1 \times A_2) \cong H_\ell(A_1) \times H_\ell(A_2)$ follows from theorem 6.4.1 and the remark following it.

Consider now the case when A_1 is of type IV. It is not hard to check (from the results in [80]) that either $\mathrm{Lie}(H_\ell(A_1)) \otimes \overline{\mathbb{Q}_\ell}$ does not have any simple factor isomorphic to $\mathfrak{sl}_2$ (cases IV(1,1) and IV(4,1) in the notation of [80]) or we are in case IV(2,1). In the former case we apply proposition 6.3.9 to deduce that $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2)$ for all primes ℓ. Suppose instead that we are in case IV(2,1), that is to say $\mathrm{End}^0_{\overline{K}}(A_1)$ is a CM field E of degree 4 over $\mathbb{Q}$. Let E_0 be the maximal totally real subfield of E. We read from [80] the equality $H(A_1)^{\mathrm{der}} = \mathrm{Res}_{E_0/\mathbb{Q}} \mathrm{SU}(E^2, \psi)$, where ψ is a suitable Hermitian form on E^2. Since $[E_0 : \mathbb{Q}] = 2$ and $\mathrm{SU}(E^2, \psi)$ is an E_0-form of SL_2, the group $H(A_1)^{\mathrm{der}}$ is a $\mathbb{Q}$-form of SL_2^2; moreover, it is $\mathbb{Q}$-simple by Theorem 1.10 of [99]. Finally, the Mumford-Tate conjecture holds for A_1 by theorem 6.2.19, so for all primes ℓ we have an isomorphism $H_\ell(A_1) \cong H(A_1) \otimes \mathbb{Q}_\ell$. By lemma 6.3.2 there is a prime p such that the group $H_p(A_1)^{\mathrm{der}} \cong H(A_1)^{\mathrm{der}} \otimes \mathbb{Q}_p$ is simple over $\mathbb{Q}_p$.
Suppose by contradiction that $\mathrm{rk}\, H_p(A)^{\mathrm{der}}$ is strictly less than $\mathrm{rk}\, H_p(A_1)^{\mathrm{der}} + \mathrm{rk}\, H_p(A_2)$. As $\mathrm{rk}\, H_p(A_2) = 1$ we have $\mathrm{rk}\, H_p(A)^{\mathrm{der}} = \mathrm{rk}\, H_p(A_1)^{\mathrm{der}}$, so the natural projection $H_p(A) \twoheadrightarrow H_p(A_1)$ induces an isogeny $H_p(A)^{\mathrm{der}} \to H_p(A_1)^{\mathrm{der}}$. Since $H_p(A_1)^{\mathrm{der}}$ is simple, the same is true for $H_p(A)^{\mathrm{der}}$; but this is absurd, because the canonical projection $H_p(A)^{\mathrm{der}} \twoheadrightarrow H_p(A_2)$ then gives a surjective morphism in which the source $H_p(A)^{\mathrm{der}}$ is simple but does not have the same rank as the image $H_p(A_2)$. The contradiction shows that $\mathrm{rk}\, H_p(A)^{\mathrm{der}} = \mathrm{rk}\, H_p(A_1)^{\mathrm{der}} + \mathrm{rk}\, H_p(A_2)$, from which we deduce first that $H_p(A) \cong H_p(A_1) \times H_p(A_2)$ and then (since the ranks of $H_\ell(A_1), H_\ell(A_2)$ and $H_\ell(A)$ do not depend on ℓ) that $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2)$ holds for all primes ℓ.

We finally come to the case $\dim A_1 = 4$ and $\mathrm{End}_{\overline{K}}(A_1) = \mathbb{Z}$. If for one (hence every) prime ℓ we have $H_\ell(A_1) = \mathrm{Sp}_{8,\mathbb{Q}_\ell}$, then the abelian variety A_1 is of general Lefschetz type (cf. [80, §4.1]), so the equality $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2)$ follows from theorem 6.4.1. Thus the last case we have to cover is that of $H_\ell(A_1)$ being a $\mathbb{Q}_\ell$-form of SL_2^3 for every prime ℓ. By [98, Theorem 5.13], there is a simple $\mathbb{Q}$-algebraic group $P(A_1)$ such that, for a set of primes ℓ of Dirichlet density 1, we have an isomorphism $H_\ell(A_1) \cong P(A_1) \otimes \mathbb{Q}_\ell$. Furthermore, $P(A_1)$ is a $\mathbb{Q}$-form of SL_2^3, so by lemma 6.3.2 we can choose a prime p for which $H_p(A_1) \cong P(A_1) \otimes \mathbb{Q}_p$ is $\mathbb{Q}_p$-simple. We can now repeat the argument of case IV(2,1) above: if by contradiction we had $\mathrm{rk}\, H_p(A) < \mathrm{rk}\, H_p(A_1) + \mathrm{rk}\, H_p(A_2)$ then $H_p(A)$ would be simple, and the canonical projection from $H_p(A)$ to $H_p(A_2)$ would be a surjective morphism between groups of different rank, which is absurd because the source is simple. We deduce once more that $\mathrm{rk}\, H_\ell(A) = \mathrm{rk}\, H_\ell(A_1) + \mathrm{rk}\, H_\ell(A_2)$ holds for $\ell = p$ (hence for every prime ℓ), so for every ℓ we have $H_\ell(A) \cong H_\ell(A_1) \times H_\ell(A_2)$. □

Chapter 7

Torsion points and roots of unity

7.1 Introduction

In this chapter we consider the following problem: given a number field K, an abelian variety A/K (of dimension g), a prime ℓ, and a finite subgroup H of $A[\ell^\infty]$, how does the number field $K(H)$ intersect the ℓ-cyclotomic extension $K(\mu_{\ell^\infty})$? More precisely, is the intersection completely accounted for by the fact that $K(H)$ contains the image of the Weil pairing $H \times H \to \mu_{\ell^\infty}$? In order to study this question, Hindry and Ratazzi have introduced in [38] and [39] two variants of a property they call (μ), and which we now recall. We fix a polarization $\varphi : A \to A^\vee$ and, for every $n \geq 0$, we denote by e_{ℓ^n} the ℓ^n-Weil pairing $A[\ell^n] \times A[\ell^n] \to \mu_{\ell^n}$ given by composing the usual Weil pairing $A[\ell^n] \times A^\vee[\ell^n] \to \mu_{\ell^n}$ with the map $A[\ell^n] \to A^\vee[\ell^n]$ induced by φ. If H is a finite subgroup of $A[\ell^\infty]$ we now define

$$m_1(H) = \max\{k \in \mathbb{N} \mid \exists n \geq 0,\ \exists P, Q \in H \text{ of order } \ell^n \text{ such that } e_{\ell^n}(P,Q) \text{ generates } \mu_{\ell^k}\}.$$

Following [39, Définition 3.8] we can then introduce the following definition:

Definition 7.1.1. We say that $(A/K, \varphi)$ satisfies property $(\mu)_s$ (where "s" stands for "strong") if there exists a constant $C > 0$, depending on A/K and φ, such that for all primes ℓ and all finite subgroups H of $A[\ell^\infty]$ the following inequalities hold:

$$\frac{1}{C}[K(\mu_{\ell^{m_1(H)}}) : K] \leq [K(H) \cap K(\mu_{\ell^\infty}) : K] \leq C[K(\mu_{\ell^{m_1(H)}}) : K].$$

Remark 7.1.2. It is easy to see that the choice of the polarization φ plays essentially no role, and $(A/K, \varphi)$ satisfies property $(\mu)_s$ for a given φ if and only $(A/K, \psi)$ satisfies property $(\mu)_s$ for every polarization ψ of A/K (possibly for different values of the constant C); for this reason we shall simply say that A/K satisfies property $(\mu)_s$ when it does for one (hence any) polarization. It is shown in [39] that if A/K satisfies the Mumford-Tate conjecture and has Mumford-Tate group isomorphic to $\mathrm{GSp}_{2\dim A,\mathbb{Q}}$, then property $(\mu)_s$ holds for A.

We also consider the following variant of property $(\mu)_s$, which we call $(\mu)_w$ ("weak"), and which was first introduced in [38, Définition 6.3]:

Definition 7.1.3. We say that A satisfies property $(\mu)_w$ if the following is true: there exists a constant $C > 0$, depending on A/K, such that for all primes ℓ and all finite subgroups H of $A[\ell^\infty]$

there exists $n \in \mathbb{N}$ (in general depending on ℓ and H) such that

$$\frac{1}{C}\,[K(\mu_{\ell^n}) : K] \leq [K(H) \cap K(\mu_{\ell^\infty}) : K] \leq C\,[K(\mu_{\ell^n}) : K]\,. \tag{7.1}$$

Clearly, property $(\mu)_s$ implies property $(\mu)_w$. In this chapter we show the following two results:

Theorem 7.1.4. *Let K be a number field and A/K be an abelian variety. If A satisfies the Mumford-Tate conjecture, then property $(\mu)_w$ holds for A.*

Theorem 7.1.5. *There exists an abelian fourfold A, defined over a number field K, such that $\mathrm{End}_{\overline{K}}(A) = \mathbb{Z}$ and for which property $(\mu)_s$ does not hold. More precisely, such an A can be taken to be any member of the family constructed by Mumford in [85].*

7.2 Property $(\mu)_w$

7.2.1 Preliminaries

We fix once and for all an embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}$, and consider the number field K as a subfield of $\overline{\mathbb{Q}} \subseteq \mathbb{C}$. The letter A denotes a fixed abelian variety over K; if ℓ is a prime number and n is a positive integer, we write G_{ℓ^n} for the Galois group of $K(A[\ell^n])/K$ and G_{ℓ^∞} for the Galois group of $K(A[\ell^\infty])/K$. Finally, we take the following definition for the Mumford-Tate group of A:

Definition 7.2.1. Let K be a number field, A/K be an abelian variety, and V be the $\mathbb{Q}$-vector space $H_1(A(\mathbb{C}), \mathbb{Q})$, equipped with its natural Hodge structure of weight -1. Also let $V_{\mathbb{Z}} = H_1(A(\mathbb{C}), \mathbb{Z})$, write $\mathbb{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_{m,\mathbb{C}})$ for Deligne's torus, and let $h : \mathbb{S} \to \mathrm{GL}_{V\otimes\mathbb{R}}$ be the morphism giving V its Hodge structure. We define $\mathrm{MT}(A)$ to be the $\mathbb{Q}$-Zariski closure of the image of h in GL_V, and extend it to a scheme over $\mathbb{Z}$ by taking its $\mathbb{Z}$-closure in $\mathrm{GL}_{V_{\mathbb{Z}}}$.

Remark 7.2.2. Taking the $\mathbb{Z}$-Zariski closure in the previous definition allows us to consider points of $\mathrm{MT}(A)$ with values in arbitrary rings. Notice that $\mathrm{MT}(A)$, being an algebraic group over a field of characteristic 0, is smooth by Cartier's theorem. It follows that $\mathrm{MT}(A)$ is smooth over an open subscheme of $\mathrm{Spec}\,\mathbb{Z}$.

The following theorem summarizes fundamental results, due variously to Serre [122], Wintenberger [145], Deligne [23, I, Proposition 6.2], Borovoĭ [14] and Pjateckiĭ-Šapiro [100], on the structure of Galois representations arising from abelian varieties over number fields; see also [40, §10] for a detailed proof of the last statement.

Theorem 7.2.3. *Let K be a number field and A/K be an abelian variety. The group $\mathrm{MT}(A)$ is smooth over an open subscheme of $\mathrm{Spec}\,\mathbb{Z}$. There exists a finite extension L of K such that for all primes ℓ the image of the natural representation $\rho_{\ell^\infty} : \mathrm{Gal}(\overline{L}/L) \to \mathrm{Aut}\,T_\ell(A)$ lands into $\mathrm{MT}(A)(\mathbb{Z}_\ell)$, and likewise the image of $\rho_\ell : \mathrm{Gal}(\overline{L}/L) \to \mathrm{Aut}\,A[\ell]$ lands into $\mathrm{MT}(A)(\mathbb{F}_\ell)$. If furthermore the Mumford-Tate conjecture holds for A, then the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : \mathrm{Im}\,\rho_{\ell^\infty}]$ is bounded by a constant independent of ℓ; the same is true for $[\mathrm{MT}(A)(\mathbb{F}_\ell) : \mathrm{Im}\,\rho_\ell]$.*

In view of the result we want to prove (theorem 7.1.4), we assume from now on that the Mumford-Tate conjecture is true for our abelian variety A. As the statement of theorem 7.1.4 is clearly invariant under extensions of the base field, theorem 7.2.3 allows us to assume that $\rho_{\ell^\infty}\left(\mathrm{Gal}(\overline{K}/K)\right)$ is included in $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ for all primes ℓ, in such a way that the index $[\mathrm{MT}(A)(\mathbb{Z}_\ell) : \mathrm{Gal}\left(K(A[\ell^\infty])/K\right)]$ is bounded by a constant independent of ℓ. Since the statement of theorem 7.1.4 is also invariant under isogenies, making a further extension of the base field if necessary we can also assume without loss of generality that A is principally polarized, which implies that G_{ℓ^∞}, resp G_ℓ, is a subgroup of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$, resp. of $\mathrm{GSp}_{2g}(\mathbb{F}_\ell)$.
The following simple lemma shows that the property of having index bounded by a constant is stable under passage to subgroups and quotients:

Lemma 7.2.4. *Let C be a group and A, B be subgroups of C such that $[C : B]$ is finite. We have $[A : B \cap A] \leq [C : B]$. Moreover, if $\pi : C \to D$ is a quotient of C, then $[D : \pi(B)] \,\big|\, [C : B]$.*

Proof. The map $A \hookrightarrow C \to C/B$ induces an injection (of sets) of $A/(A \cap B)$ into C/B. The second statement is obvious. □

The previous lemma allows us to work with "equalities up to a finite index", for which we now introduce some notations. If L_1, L_2 are number fields that depend on A/K and on some other set of parameters, we write $L_1 \doteq L_2$ to mean that there exists a constant C (depending on A/K only) such that the inequalities $[L_1 : L_1 \cap L_2] \leq C$ and $[L_2 : L_1 \cap L_2] \leq C$ hold for all values of the parameters; likewise, if G_1, G_2 are subgroups of a same group (and depend on some set of parameters), we write $G_1 \doteq G_2$ if both $[G_1 : G_1 \cap G_2]$ and $[G_2 : G_1 \cap G_2]$ are bounded by a constant depending only on A/K, uniformly in all other parameters. Furthermore, for two functions $f, g : I \to \mathbb{R}^+$, where I is any set, we write $f \doteq g$ if there is a constant $C' > 0$ such that $\frac{1}{C'} g(x) \leq f(x) \leq C' g(x)$ for all $x \in I$. Finally, to deal with arithmetic functions we introduce the following definition:

Definition 7.2.5. Let $\mathcal{P}$ be the set of prime numbers, I be any set and $h : I \times \mathcal{P} \to \mathbb{N}^+$ be any function. We say that $h(x, \ell)$ is a power of ℓ up to a bounded constant if there exists a $C'' > 0$ such that for all $x \in I$ and $\ell \in \mathcal{P}$ we have $\frac{h(x,\ell)}{\ell^{v_\ell(h(x,\ell))}} \leq C''$, or equivalently, if the prime-to-ℓ part of $h(x, \ell)$ is bounded independently of x and ℓ.

As a typical example of the use of this notation, notice that our assumption that we are in the situation of theorem 7.2.3 can be expressed by writing $\mathrm{Gal}\left(K(A[\ell^\infty])/K\right) \doteq \mathrm{MT}(A)(\mathbb{Z}_\ell)$ and $\mathrm{Gal}\left(K(A[\ell])/K\right) \doteq \mathrm{MT}(A)(\mathbb{F}_\ell)$. We can also apply lemma 7.2.4 to the groups $C = \mathrm{MT}(A)(\mathbb{F}_\ell)$, $B = \mathrm{Gal}\left(K(A[\ell])/K\right)$ and $A = \left\{x \in \mathrm{MT}(A)(\mathbb{F}_\ell) \,\middle|\, xh = h \quad \forall h \in H\right\}$ to get

$$\mathrm{Gal}\left(K(A[\ell])/K(H)\right) \doteq \left\{x \in \mathrm{MT}(A)(\mathbb{F}_\ell) \,\middle|\, xh = h \quad \forall h \in H\right\},$$

where the implied constant depends on A/K, but not on ℓ or H. Finally, notice that if A, B are groups (depending on some set of parameters) such that $[B : A] \leq N$ for all values of the parameters, then taking $N' := N!$ we have $[B : A] \,\big|\, N'$, again for any choice of the parameters: if we so desire we can therefore replace boundedness conditions by divisibility conditions.

7.2.2 Stabilizers of finite subgroups of $A[\ell^\infty]$

Denote by $\mathcal{G}$ the $\mathbb{Z}_\ell$-algebraic group $\mathrm{MT}(A) \times_{\mathbb{Z}} \mathbb{Z}_\ell$ and let H be a finite subgroup of $A[\ell^\infty]$. Write H as $\prod_{i=1}^{2g} \mathbb{Z}/\ell^{m_i}\mathbb{Z}$ for certain integers $m_1 \geq \ldots \geq m_{2g}$, let $e_1, \ldots, e_{2g}$ be generators of H (so e_i is a torsion point of order ℓ^{m_i}), and let $\widehat{e}_1, \ldots, \widehat{e}_{2g}$ be a basis of $T_\ell A$ lifting the e_i's (that is, satisfying $\widehat{e}_i \equiv e_i \pmod{\ell^{m_i}}$ for $i = 1, \ldots, 2g$). For a subset I of $\{1, \ldots, 2g\}$ we let $\mathcal{G}_I$ be the $\mathbb{Z}_\ell$-algebraic group given by

$$\mathcal{G}_I = \left\{ M \in \mathcal{G} \;\middle|\; M\widehat{e}_i = \widehat{e}_i \quad \forall i \in I \right\}.$$

We plan to show that $\mathcal{G}_I$ and other related groups are smooth (over $\mathbb{Z}_\ell$, or equivalently over $\mathbb{F}_\ell$, cf. lemma 7.2.9) whenever ℓ is sufficiently large with respect to A/K, independently of the choice of $\widehat{e_1}, \ldots, \widehat{e_{2g}}$ and I (the result crucial to our applications is lemma 7.2.10). We shall make repeated use of the following fact:

Theorem 7.2.6. *Let ℓ be a prime number and k be a finite field of characteristic ℓ. Let $\mathcal{F}$ be an affine group scheme over k with coordinate ring R. The following are equivalent:*

1. *$\mathcal{F}$ is smooth;*
2. *$R \otimes_k \overline{k}$ is reduced;*
3. *the nilpotency index of $R \otimes_k \overline{k}$ is smaller than ℓ, that is, there exists an integer $e < \ell$ such that for all $a \in R \otimes_k \overline{k}$ and all positive integers n, the equality $a^n = 0$ implies $a^e = 0$;*
4. *the equality $\dim_k \mathrm{Lie}\, \mathcal{F} = \dim \mathcal{F}$ holds.*

Proof. 1 and 2 are equivalent by [140, Theorem on p. 88]. 1 and 4 are equivalent by [140, Corollary on p. 94]. Clearly 2 implies 3, and 3 implies 2 by the same argument that proves Cartier's theorem (all algebraic groups over a field of characteristic zero are smooth), see for example [77, Proof of Theorem 10.1]. □

The following proposition, while certainly well-known to experts, does not seem to appear anywhere in the literature; we will use it as a substitute for Cartier's theorem on smoothness when working over a field of positive characteristic.

Proposition 7.2.7. *Let n, d, m be fixed positive integers. There is a constant $c(n,d,m)$ with the following property: for every prime $\ell > c(n,d,m)$, every finite field k of characteristic ℓ, and every algebraic subgroup $\mathcal{F}$ of $\mathrm{GL}_{n,k}$ that is cut in $\dfrac{k[x_{ij}, y]}{(\det(x_{ij})y - 1)}$ by at most m equations of degree at most d is smooth over k.*

Proof. Let $I = (f_1, \ldots, f_t)$ be the ideal defining $\mathcal{F}$ in $R := \dfrac{k[x_{ij}, y]}{(\det(x_{ij})y - 1)}$, where $t \leq m$ and the total degree of every f_h is at most d. To test smoothness we can base-change to $\overline{k}$, and by theorem 7.2.6 we only need to prove that the nilpotency index of

$$R \otimes_k \overline{k} \cong \frac{\overline{k}[x_{ij}, y]}{(\det(x_{ij})y - 1, f_1, \ldots, f_t)}$$

is bounded by a function of n, d and m alone, uniformly in ℓ and k. Now just notice that the ideal $(\det(x_{ij})y - 1, f_1, \ldots, f_t)$ is generated by equations whose number and degree are bounded in terms

of n, d, and m, so the result follows from [47, Theorem 1.3] (see also [51]). More precisely, since we have at most $m+1$ equations of degree at most $\max\{d, n+1\}$, [47, Theorem 1.3] shows that one can take $c(n,d,m) = \max\{d, n+1\}^{m+1}$. □

Lemma 7.2.8. *Let n be a positive integer, $\mathcal{F}$ be a group subscheme of $\mathrm{GL}_{n,\mathbb{Q}_\ell}$, and let $\underline{\mathcal{F}}$ be the Zariski closure of $\mathcal{F}$ in $\mathrm{GL}_{n,\mathbb{Z}_\ell}$. Then $\underline{\mathcal{F}}$ is flat over $\operatorname{Spec}\mathbb{Z}_\ell$.*

Proof. An affine scheme $\operatorname{Spec} R$ over $\mathbb{Z}_\ell$ is flat if and only if its coordinate ring R is a torsion-free $\mathbb{Z}_\ell$-module ([64, Corollary 2.14]). In our case, if I is the ideal of $\frac{\mathbb{Q}_\ell[x_{ij},y]}{(\det(x_{ij})y-1)}$ that defines $\mathcal{F}$, then $\underline{I} := I \cap \frac{\mathbb{Z}_\ell[x_{ij},y]}{(\det(x_{ij})y-1)}$ is the ideal defining $\underline{\mathcal{F}}$. In particular, the coordinate ring $\underline{R}$ of $\underline{\mathcal{F}}$ injects into the coordinate ring R of $\mathcal{F}$, which is torsion-free since it is a $\mathbb{Q}_\ell$-vector space. □

Lemma 7.2.9. *Let n be a positive integer, $\mathcal{F}$ be a group subscheme of $\mathrm{GL}_{n,\mathbb{Q}_\ell}$, and let $\underline{\mathcal{F}}$ be the Zariski closure of $\mathcal{F}$ in $\mathrm{GL}_{n,\mathbb{Z}_\ell}$. Suppose furthermore that $\underline{\mathcal{F}}$ is smooth over $\mathbb{F}_\ell$: then $\underline{\mathcal{F}}$ is smooth over $\mathbb{Z}_\ell$.*

Proof. In order for a scheme $\underline{\mathcal{F}} \,/\, \operatorname{Spec}\mathbb{Z}_\ell$ to be smooth, it is necessary and sufficient that it is locally finitely presented and flat, with fibers that are smooth varieties all of the same dimension. Finite presentation is obvious in our context, and flatness follows from the previous lemma. The dimension of the fibers is locally constant by flatness, hence constant since the only open subset of $\operatorname{Spec}\mathbb{Z}_\ell$ containing the closed point is all of $\operatorname{Spec}\mathbb{Z}_\ell$. It remains to show smoothness of the fibers: the generic fiber is smooth by Cartier's theorem ([140, §11.4]), and the special fiber is smooth by assumption. □

We finally come to the central result of this section:

Lemma 7.2.10. *For all ℓ sufficiently large (depending only on A/K), for all $\mathbb{Z}_\ell$-bases $\widehat{e}_1, \dots, \widehat{e}_{2g}$ of $T_\ell A$, and for all subsets I of $\{1, \dots, 2g\}$, the stabilizer $\mathcal{G}_I$ in $\mathcal{G}$ of the vectors $\widehat{e}_i$ (for $i \in I$) is smooth over $\mathbb{Z}_\ell$. Suppose furthermore that A/K is principally polarized, so that $\mathrm{MT}(A) \subseteq \mathrm{GSp}_{2g,\mathbb{Z}}$. Let furthermore $\lambda : \mathrm{MT}(A) \to \mathbb{G}_{m,\mathbb{Z}}$ be the restriction to $\mathrm{MT}(A)$ of the algebraic multiplier character $\lambda : \mathrm{GSp}_{2g,\mathbb{Z}} \to \mathbb{G}_{m,\mathbb{Z}}$. For ℓ large enough (again depending only on A/K) all the groups*

$$\mathcal{G}_I^{(1)} = \left\{ M \in \mathcal{G} \;\middle|\; M \cdot \widehat{e}_i = \widehat{e}_i \; \forall i \in I, \lambda(M) = 1 \right\}$$

are smooth.

Proof. Notice first that $\mathcal{G}_I$ can be obtained as the $\mathbb{Z}_\ell$-Zariski closure of the $\mathbb{Q}_\ell$-group scheme

$$\left\{ M \in H_\ell \;\middle|\; M\widehat{e}_i = \widehat{e}_i, \; \forall i \in I \right\}.$$

By lemma 7.2.9 it then suffices to prove smoothness over $\mathbb{F}_\ell$, and to do this we can base-change to $\overline{\mathbb{F}_\ell}$. We can also assume that $\mathcal{G}$ is smooth over $\mathbb{Z}_\ell$, since this is true except for finitely many exceptions. Now we claim that $\mathcal{G}_I$ is defined by equations whose number and degree are independent of ℓ: indeed, they are the equations defining $\mathrm{MT}(A)$ (and these do not depend on ℓ), together with linear equations that express in coordinates the equalities $M\widehat{e}_i = \widehat{e}_i$. Since there are at most $(2g)^2$ such linear equations, the claim follows. We then deduce from proposition 7.2.7 that for ℓ large enough $(\mathcal{G}_I)_{\mathbb{F}_\ell}$ is smooth, and an entirely similar argument also proves the result for $\mathcal{G}_I^{(1)}$. □

7.2.3 Some Galois cohomology

Lemma 7.2.11. *Let $\mathcal{G}$ be a finite étale group scheme of order N over $\mathbb{F}_\ell$. The first cohomology group $H^1(\mathbb{F}_\ell, \mathcal{G})$ is finite, of order not exceeding N.*

Proof. Recall ([140, §6.4]) that the association $\mathcal{G} \mapsto \mathcal{G}(\overline{\mathbb{F}_\ell})$ establishes an equivalence between the category of étale group schemes over $\mathbb{F}_\ell$ and that of finite groups with a continuous action of $\mathrm{Gal}\left(\overline{\mathbb{F}_\ell}/\mathbb{F}_\ell\right)$. To prove the lemma it is thus enough to consider the cohomology $H^1(\mathbb{F}_\ell, G)$ of a finite group G of order N equipped with a continuous action of $\hat{\mathbb{Z}} \cong \mathrm{Gal}\left(\overline{\mathbb{F}_\ell}/\mathbb{F}_\ell\right)$. An element of $H^1\left(\hat{\mathbb{Z}}, G\right)$ is represented by a continuous map $\hat{\mathbb{Z}} \to G$, which in turn is uniquely determined by the image of a topological generator of $\hat{\mathbb{Z}}$: it follows that there are no more than $|G| = N$ such maps, hence that the order of $H^1(\hat{\mathbb{Z}}, G)$ is bounded by N as claimed. □

Lemma 7.2.12. *Let $\mathcal{G}$ be a linear algebraic group over $\mathbb{F}_\ell$. We have $|H^1(\mathbb{F}_\ell, \mathcal{G})| \leq |H^1(\mathbb{F}_\ell, \mathcal{G}/\mathcal{G}^0)|$, so in particular the order of $H^1(\mathbb{F}_\ell, \mathcal{G})$ does not exceed the order of the group of components of $\mathcal{G}$.*

Proof. The long exact sequence in cohomology associated with the sequence

$$1 \to \mathcal{G}^0 \to \mathcal{G} \to \mathcal{G}/\mathcal{G}^0 \to 1$$

contains in particular the segment $H^1(\mathbb{F}_\ell, \mathcal{G}^0) \to H^1(\mathbb{F}_\ell, \mathcal{G}) \to H^1(\mathbb{F}_\ell, \mathcal{G}/\mathcal{G}^0)$, where the first term is trivial by Lang's theorem (any connected algebraic group over a finite field has trivial H^1, [54, Theorem 2]). The first statement follows. The second is then a consequence of the previous lemma and of the fact that $\mathcal{G}/\mathcal{G}^0$ is étale by [140, §6.7]. □

7.2.4 Proof of theorem 7.1.4

We now come to the core of the proof of theorem 7.1.4.

Lemma 7.2.13. *Suppose the Mumford-Tate conjecture holds for A: then for all primes ℓ and for all finite subgroups H of $A[\ell]$ there exists $m \in \{0, 1\}$ such that*

$$[K(\mu_{\ell^m}) : K] \stackrel{\circ}{=} [K(H) \cap K(\mu_\ell) : K], \tag{7.2}$$

that is to say, there exists $D > 0$ (depending on A/K) with the following property: for every ℓ and every subgroup H of $A[\ell]$ there exists $m \in \{0, 1\}$ such that

$$D^{-1}\,[K(H) \cap K(\mu_\ell) : K] \leq [K(\mu_{\ell^m}) : K] \leq D\,[K(H) \cap K(\mu_\ell) : K]. \tag{7.3}$$

Proof. Recall that we denote λ the restriction to $\mathrm{MT}(A)$ of the multiplier map $\mathrm{GSp}_{2g,\mathbb{Z}} \to \mathbb{G}_{m,\mathbb{Z}}$. Observe first that it suffices to prove that the conclusion of the lemma holds for all but finitely many primes: indeed, for a fixed prime ℓ the finite group $A[\ell]$ possesses only finitely many subgroups H, so we can choose D so large that (7.3) holds for any such H (with $m = 0$, say). Disregarding a finite set of primes (which we call "bad") we can therefore assume that ℓ is large enough that the groups $\mathcal{G}_I$ and $\mathcal{G}_I^{(1)}$ of section 7.2.2 are smooth (lemma 7.2.10) and that ℓ is unramified in K. Recall that $G_\ell \subseteq \mathrm{GSp}_{2g}(\mathbb{F}_\ell)$ is a subgroup of $\mathrm{MT}(A)(\mathbb{F}_\ell)$, and that (since we assume A to be principally polarized) for all primes ℓ we have $\lambda \circ \rho_\ell = \chi_\ell$, the mod-$\ell$ cyclotomic character. Let now $e_1, \ldots, e_{2g}$ be an $\mathbb{F}_\ell$-basis of $A[\ell]$ such that $e_1, \ldots, e_r$ is an $\mathbb{F}_\ell$-basis of H. We consider the finite

group $S = \left\{M \in G_\ell \;\middle|\; M \cdot h = h \quad \forall h \in H\right\}$, that is, the stabilizer of H in G_ℓ, and the algebraic group $\mathcal{S} = \left\{M \in \mathrm{MT}(A)_{\mathbb{F}_\ell} \;\middle|\; M \cdot e_i = e_i,\ 1 \le i \le r\right\}$, that is, the stabilizer of H in $\mathrm{MT}(A)_{\mathbb{F}_\ell}$.
The group $\mathcal{S}$ is obtained as the special fiber of a group of the form $\mathcal{G}_I$ (notation as in section 7.2.2), hence is smooth by our assumption on ℓ. It is clear by definition that $S = G_\ell \cap \mathcal{S}(\mathbb{F}_\ell)$; since $G_\ell \doteq \mathrm{MT}(A)(\mathbb{F}_\ell)$, this shows in particular that $S \doteq \mathcal{S}(\mathbb{F}_\ell)$. We now claim that the group of components of $\mathcal{S}$ has order bounded by a constant B independent of ℓ and H. Notice first that it is enough to bound the number of $\overline{\mathbb{F}_\ell}$-points of the group of components of $\mathcal{S}$, hence it is enough to consider the number of irreducible components of $\mathcal{S}_{\overline{\mathbb{F}_\ell}}$. Next observe that $\mathcal{S}_{\overline{\mathbb{F}_\ell}}$ is cut in $M_{2g}(\overline{\mathbb{F}_\ell}) \cong \mathbb{A}^{(2g)^2}_{\overline{\mathbb{F}_\ell}}$ by the equations defining $\mathrm{MT}(A)$ (and these only depend on A/K) and by the $2g \cdot r$ hyperplanes given by the vector equations $M \cdot e_i = e_i$ for $i = 1, \dots, r$: since clearly $r \le 2g$, we see that $\mathcal{S}_{\overline{\mathbb{F}_\ell}}$ is defined by equations whose number and degree are bounded uniformly in ℓ and H. By a variant of Bézout's theorem (see [139, Theorem 7.1] for a precise statement), this implies that the number of irreducible components of $\mathcal{S}_{\overline{\mathbb{F}_\ell}}$ is bounded uniformly in ℓ and H. The very same argument also shows that the order of the group of connected components of

$$\mathcal{S}_1 = \left\{M \in \mathrm{MT}(A)_{\mathbb{F}_\ell} \;\middle|\; M \cdot h = h \ \ \forall h \in H,\ \ \lambda(M) = 1\right\} = \ker(\lambda : \mathcal{S} \to \mathbb{G}_{m,\mathbb{F}_\ell})$$

is bounded by a constant independent of ℓ, which we call B_1. Notice furthermore that the group $S_1 = \left\{M \in G_\ell \;\middle|\; M \cdot h = h \quad \forall h \in H,\ \lambda(M) = 1\right\}$ satisfies $S_1 \doteq \mathcal{S}_1(\mathbb{F}_\ell)$.

Consider now the restriction of $\lambda : \mathrm{GSp}_{2g,\mathbb{F}_\ell} \to \mathbb{G}_{m,\mathbb{F}_\ell}$ to $\mathcal{S}^0$, the identity component of $\mathcal{S}$. Notice that (for the ℓ we are considering) the group $\mathcal{S}^0$ is smooth (lemma 7.2.10), so the image $\lambda(\mathcal{S}^0)$ is a connected reduced subgroup of $\mathbb{G}_{m,\mathbb{F}_\ell}$, hence it is either trivial or all of $\mathbb{G}_{m,\mathbb{F}_\ell}$. Let us consider the two cases separately.

$\lambda(\mathcal{S}^0)$ is trivial. As we have already remarked we have $S \subseteq \mathcal{S}(\mathbb{F}_\ell)$. It follows that the order of $\lambda(S)$ is at most the order of $\lambda(\mathcal{S}(\mathbb{F}_\ell))$, which in turn does not exceed $[\mathcal{S} : \mathcal{S}^0]$ since the restriction of λ to $\mathcal{S}^0$ is trivial. Hence we have $|\lambda(S)| \le [\mathcal{S} : \mathcal{S}^0] \le B$.

$\lambda : \mathcal{S}^0 \to \mathbb{G}_{m,\mathbb{F}_\ell}$ is onto. Consider the exact sequence

$$1 \to \mathcal{S}_1 \to \mathcal{S} \xrightarrow{\lambda} \mathbb{G}_{m,\mathbb{F}_\ell} \to 1$$

and take $\mathbb{F}_\ell$-rational points: the associated long exact sequence in cohomology shows that

$$\mathcal{S}(\mathbb{F}_\ell) \xrightarrow{\lambda} \mathbb{G}_{m,\mathbb{F}_\ell}(\mathbb{F}_\ell) = \mathbb{F}_\ell^\times \to H^1\left(\mathbb{F}_\ell, \mathcal{S}_1\right)$$

is exact, so $\left|\mathrm{coker}\left(\mathcal{S}(\mathbb{F}_\ell) \xrightarrow{\lambda} \mathbb{F}_\ell^\times\right)\right|$ is at most $\left|H^1\left(\mathbb{F}_\ell, \mathcal{S}_1\right)\right|$, which in turn (by lemma 7.2.12 and what we have already proved) does not exceed B_1. Since $S \doteq \mathcal{S}(\mathbb{F}_\ell)$, it follows that

$$|\lambda(S)| \doteq |\lambda(\mathcal{S}(\mathbb{F}_\ell))| \ge \frac{\ell - 1}{B_1},$$

that is, there exists a constant B' (independent of ℓ) such that whenever $\lambda : \mathcal{S}^0 \to \mathbb{G}_{m,\mathbb{F}_\ell}$ is onto the inequality $|\lambda(S)| \ge \dfrac{\ell - 1}{B'}$ holds. Let now B'' be a constant large enough that inequality (7.3) in the statement of the lemma holds, with $D = B''$, for all the (finitely many) bad primes ℓ, and for the (finitely many) subgroups H of $A[\ell]$ for each of these primes. Finally set $D = \max\{B, B', B''\}$. We now show that inequality (7.3) is satisfied for all primes ℓ and all subgroups H of $A[\ell]$. It is clear by construction that this is true for the bad primes, so we can again suppose that ℓ is unramified in K and that the groups $\mathcal{S}, \mathcal{S}_1$ are smooth over $\mathbb{F}_\ell$; let once more H be a subgroup of $A[\ell]$ for such an ℓ.

Observe that the group S we considered above is by definition the Galois group of $K(A[\ell])/K(H)$, whereas the Galois group of $K(A[\ell])$ over $K(\mu_\ell)$ is $N := \ker\left(G_\ell \xrightarrow{\lambda} \mathbb{F}_\ell^\times\right)$. It follows that the Galois group of $K(A[\ell])$ over $K(H) \cap K(\mu_\ell)$ is the group generated by S and N, hence the degree of $K(H) \cap K(\mu_\ell)$ over K is the index of NS in G_ℓ. On the other hand we have $|G_\ell/NS| = \frac{|G_\ell/N|}{|NS/N|}$ (recall that N is normal in G_ℓ by construction), and G_ℓ/N is clearly isomorphic to the image of $\lambda : G_\ell \to \mathbb{F}_\ell^\times$. As ℓ is unramified in K, the mod-ℓ cyclotomic character $\chi_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathbb{F}_\ell^\times$ is surjective, hence we have $\lambda(G_\ell) = \chi_\ell(\mathrm{Gal}(\overline{K}/K)) = \mathbb{F}_\ell^\times$ and therefore

$$[K(H) \cap K(\mu_\ell) : K] = |G_\ell/NS| = \frac{|\lambda(G_\ell)|}{|\lambda(NS)|} = \frac{\ell - 1}{|\lambda(S)|}.$$

By our previous arguments we now see that

- either $\lambda(\mathcal{S}^0)$ is trivial, in which case $1 \leq |\lambda(S)| \leq B$ and (7.3) is satisfied by taking $m = 1$;
- or $\lambda : \mathcal{S}^0 \to \mathbb{G}_{m,\mathbb{F}_\ell}$ is onto, in which case we have $\dfrac{\ell - 1}{B'} \leq |\lambda(S)| \leq \ell - 1$ and (7.3) is satisfied by taking $m = 0$.

□

To complete the proof of theorem 7.1.4 we need two more lemmas.

Lemma 7.2.14. *Let K be a number field and A/K be an abelian variety satisfying the Mumford-Tate conjecture: then for any finite subgroup H of $A[\ell^\infty]$ the degree $[K(H) : K(H[\ell])]$ is a power of ℓ (up to a bounded constant).*

Proof. We use the notation from section 7.2.2; in particular we write $H \cong \prod_{i=1}^{2g} \mathbb{Z}/\ell^{m_i}\mathbb{Z}$, fix generators $e_1, \ldots, e_{2g}$ of H and a basis $\widehat{e}_1, \ldots, \widehat{e}_{2g}$ of $T_\ell A$ lifting the e_i's. We suppose first that $\mathcal{G} := \mathrm{MT}(A)_{\mathbb{Z}_\ell}$ is smooth over $\mathbb{Z}_\ell$. Inspired by the approach of [39], given $\mathbb{Z}_\ell$-algebraic subgroups $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \cdots \subseteq \mathcal{G}_t$ of $\mathcal{G}$, a strictly increasing sequence $n_1 < n_2 < \cdots < n_t$ of positive integers, and a positive integer n, we now denote by $\mathcal{G}(n; n_1, \ldots, n_t)$ the finite group

$$\left\{ M \in \mathcal{G}(\mathbb{Z}/\ell^n\mathbb{Z}) \;\middle|\; M \in \mathcal{G}_i \bmod \ell^{\min(n,n_i)}, \quad i = 1, \ldots, t \right\}.$$

It is natural to include the case of t being 0: if n_i is the empty sequence, we simply define $\mathcal{G}(n)$ to be $\mathcal{G}(\mathbb{Z}/\ell^n\mathbb{Z})$. To the group H we now attach a strictly decreasing sequence of positive integers $m^{(1)} > m^{(2)} > \cdots > m^{(t)} \geq 1$ (where $t \leq 2g$) by setting

$$m^{(1)} = \max\left\{m_i \;\middle|\; m_i \neq 0\right\} \text{ and recursively } m^{(r+1)} = \max\left\{m_i \;\middle|\; 0 < m_i < m^{(r)}\right\},$$

and, for $1 \leq r \leq t$, we let $I_r = \left\{i \in \{1, \ldots, 2g\} \;\middle|\; m_i \geq m^{(r)}\right\}$.
Finally, for $1 \leq r \leq t$, we define $\mathcal{G}_r := \mathcal{G}_{I_{t+1-r}}$ (notation as in section 7.2.2) and consider the strictly *increasing* sequence $n_r = m^{(t+1-r)}$ (for $1 \leq r \leq t$). We can assume that ℓ is so large that all the groups $\mathcal{G}_r$ are smooth over $\mathbb{Z}_\ell$ (lemma 7.2.10), and, as in [39], we see that the $\mathcal{G}_r$'s so defined form an increasing sequence of subgroups of $\mathcal{G}$ such that $[K(H[\ell^m]) : K] \doteq [\mathcal{G}(\mathbb{Z}/\ell^m\mathbb{Z}) : \mathcal{G}(m; n_1, \ldots, n_t)]$. We now show that (for any H and any $m \geq 1$) the number

$$\frac{[\mathcal{G}(\mathbb{Z}/\ell^m\mathbb{Z}) : \mathcal{G}(m; n_1, \ldots, n_t)]}{[\mathcal{G}(\mathbb{Z}/\ell\mathbb{Z}) : \mathcal{G}(1; n_1, \ldots, n_t)]} \tag{7.4}$$

is a power of ℓ. To prove this fact, we preliminarily show that for all $m \geq 2$ the reduction map $\mathcal{G}\left(\mathbb{Z}/\ell^m\mathbb{Z}\right) \xrightarrow{\pi_{m-1}} \mathcal{G}\left(\mathbb{Z}/\ell^{m-1}\mathbb{Z}\right)$ maps $\mathcal{G}(m; n_1, \ldots, n_t)$ surjectively onto $\mathcal{G}(m-1; n_1, \ldots, n_t)$. We can proceed by induction on t, showing the stronger statement that this is true for any chain of groups $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \cdots \subset \mathcal{G}_t \subset \mathcal{G}$ where each term is smooth over $\mathbb{F}_\ell$. Indeed,

- for $t = 0$ the claim follows from the smoothness of $\mathcal{G}$ and Hensel's lemma;
- if $m \leq n_t$, then we have $\mathcal{G}(j; n_1, \ldots, n_t) = \mathcal{G}_t(j; n_1, \ldots, n_{t-1})$ both for $j = m$ and $j = m-1$, so the claim follows from the induction hypothesis;
- if $m > n_t$, then $\mathcal{G}\left(\mathbb{Z}/\ell^m\mathbb{Z}\right) \to \mathcal{G}\left(\mathbb{Z}/\ell^{m-1}\mathbb{Z}\right)$ is surjective by smoothness of $\mathcal{G}$, and furthermore, since by assumption we have $m - 1 \geq n_t > n_{t-1} > \ldots > n_1$, any lift to $\mathcal{G}\left(\mathbb{Z}/\ell^m\mathbb{Z}\right)$ of a point in $\mathcal{G}(m-1; n_1, \ldots, n_t)$ belongs to $\mathcal{G}(m; n_1, \ldots, n_t)$. In particular, the induced map $\mathcal{G}(m; n_1, \ldots, n_t) \to \mathcal{G}(m-1; n_1, \ldots, n_t)$ is indeed surjective.

We now prove our claim that (7.4) is a power of ℓ by induction on m. Notice that, by Hensel's lemma and since $m \geq 2$, the kernel of π_{m-1} is an ℓ-group (of order $\ell^{\dim \mathcal{G}}$). It follows that π_{m-1} induces a surjective map $\mathcal{G}(m; n_1, \ldots, n_t) \to \mathcal{G}(m-1; n_1, \ldots, n_t)$ whose kernel is an ℓ-group; in particular, the numbers $\frac{|\mathcal{G}(m; n_1, \ldots, n_t)|}{|\mathcal{G}(m-1; n_1, \ldots, n_t)|}$ and $\frac{|\mathcal{G}(\mathbb{Z}/\ell^m\mathbb{Z})|}{|\mathcal{G}(\mathbb{Z}/\ell^{m-1}\mathbb{Z})|}$ are both powers of ℓ, and an immediate induction shows that the same is true for (7.4). Choosing m large enough that $H = H[\ell^m]$, it follows from our previous considerations that $\dfrac{[K(H[\ell^m]) : K]}{[K(H[\ell]) : K]} = [K(H) : K(H[\ell])]$ is a power of ℓ (up to bounded constants), which finishes the proof of the lemma when $\mathrm{MT}(A)$ is smooth over $\mathbb{F}_\ell$, and leaves us with only the finitely many bad reduction primes to consider. To establish the lemma we thus need to show that, for ℓ ranging over the bad primes and H ranging over the finite subgroups of $A[\ell^\infty]$, the degree $[K(H) : K(H[\ell])]$ is within a constant factor of a power of ℓ. As we are only considering finitely many primes, there are only finitely many subgroups of $A[\ell]$, and therefore we have $[K(H[\ell]) : K] \stackrel{\circ}{=} 1$; hence we just need to show that $[K(H) : K]$ is a power of ℓ up to a constant factor. Let ℓ^m be the exponent of H. Since the prime-to-ℓ part of $[K(H) : K]$ divides the prime-to-ℓ part of $[K(A[\ell^m]) : K]$, it is enough to show that $|G_{\ell^m}| = |\operatorname{Gal}(K(A[\ell^m])/K)|$ is a power of ℓ up to a bounded constant. Let C be the least common multiple of the orders of the groups G_ℓ for ℓ ranging over the primes of bad reduction. Consider the reduction map $\pi : G_{\ell^m} \to G_\ell$, and notice that its kernel is a subgroup of $\ker\left(\mathrm{GL}_{2g}(\mathbb{Z}/\ell^m\mathbb{Z}) \to \mathrm{GL}_{2g}(\mathbb{F}_\ell)\right)$, hence in particular an ℓ-group; we can then write $\frac{|G_{\ell^m}|}{|\ker \pi|}$ as $|\pi(G_{\ell^m})|$, which by construction is an integer dividing C. Since $|\ker \pi|$ is a power of ℓ, we see that the prime-to-ℓ part of $|G_{\ell^m}|$ is bounded by C; this completes the proof in the non-smooth case as well. □

Lemma 7.2.15. *Let K be a number field, A/K an abelian variety satisfying the Mumford-Tate conjecture, ℓ a prime number, and H a finite subgroup of $A[\ell^\infty]$. We have*

$$K(H) \cap K(\mu_\ell) \stackrel{\circ}{=} K(H[\ell]) \cap K(\mu_\ell),$$

and the degree of $K(H) \cap K(\mu_{\ell^\infty})$ over $K(H) \cap K(\mu_\ell)$ is a power of ℓ.

Proof. Let m be such that $H \subseteq A[\ell^m]$. The Galois group of $K(A[\ell^m])$ over $K(H[\ell]) \cap K(\mu_\ell)$ is generated by the Galois groups of $K(A[\ell^m])$ over $K(H[\ell])$ (which we denote by S_1) and over $K(\mu_\ell)$

(denoted by N); notice that $N = \ker\left(G_{\ell^m} \xrightarrow{\lambda} \mathbb{F}_\ell^\times\right)$. Let now S_m be the Galois group of $K(A[\ell^m])$ over $K(H)$. By lemma 7.2.14 we see that $[S_1 : S_m]$ is a power of ℓ (up to a constant bounded independently of ℓ), hence $[NS_1 : NS_m] = \dfrac{|NS_1/N|}{|NS_m/N|} = \dfrac{|\lambda(S_1)|}{|\lambda(S_m)|}$ is again a power of ℓ (up to a constant independent of ℓ). On the other hand, $\lambda(S_1)$ is a subgroup of $\mathbb{F}_\ell^\times$, hence of order prime to ℓ: it follows that $\left|\frac{\lambda(S_1)}{\lambda(S_m)}\right| \doteq 1$, and therefore $NS_1 \doteq NS_m$. Now NS_1 is the Galois group of $K(A[\ell^m])$ over $K(H[\ell]) \cap K(\mu_\ell)$, while NS_m is the Galois group of $K(A[\ell^m])/K(H) \cap K(\mu_\ell)$: by Galois theory, this implies $K(H) \cap K(\mu_\ell) \doteq K(H[\ell]) \cap K(\mu_\ell)$ as claimed. The second part is immediate by Galois theory. □

Theorem 7.2.16. *(Theorem 7.1.4) Let K be a number field and A/K be an abelian variety. If A satisfies the Mumford-Tate conjecture, then property $(\mu)_w$ holds for A.*

Proof. Fix a prime ℓ and a subgroup $H \subseteq A[\ell^\infty]$: we want to show that we can choose n so as to satisfy inequality (7.1) (for some constant C only depending on A/K). Let L be the intersection $K(H[\ell]) \cap K(\mu_\ell)$. By lemma 7.2.13, we can choose $m \in \{0, 1\}$ so that

$$[L : K] \doteq [K(\mu_{\ell^m}) : K], \tag{7.5}$$

and by lemma 7.2.15 we see that there is an integer j such that $[K(H) \cap K(\mu_{\ell^\infty}) : L] \doteq \ell^j$. Observe now that $[K(H) \cap K(\mu_{\ell^\infty}) : K] = [K(H) \cap K(\mu_{\ell^\infty}) : L]\,[L : K] \doteq \ell^j [L : K]$, hence by (7.5) we have $[K(H) \cap K(\mu_{\ell^\infty}) : K] \doteq \ell^j \cdot [K(\mu_{\ell^m}) : K]$. Using the obvious equalities (up to bounded constants) $[K(\mu_{\ell^{j+1}}) : K(\mu_\ell)] \doteq [K(\mu_{\ell^j}) : K] \doteq \ell^j$ we deduce

$$\begin{aligned}
[K(H) \cap K(\mu_{\ell^\infty}) : K] &\doteq \ell^j \cdot [K(\mu_{\ell^m}) : K] \\
&\doteq [K(\mu_{\ell^{j+m}}) : K(\mu_{\ell^m})] \cdot [K(\mu_{\ell^m}) : K] \\
&= [K(\mu_{\ell^{j+m}}) : K].
\end{aligned}$$

This shows that, if we take C to be the constant implied in the last formula, for all primes ℓ and all finite subgroups H of $A[\ell^\infty]$ inequality (7.1) can be satisfied by taking $n = m + j$, and therefore property $(\mu)_w$ holds for A as claimed. □

7.3 Property $(\mu)_s$

Let F be any field. We start by considering the representation

$$\begin{array}{rccc}
\rho : & \mathrm{GL}_2(F) \times \mathrm{GL}_2(F) \times \mathrm{GL}_2(F) & \to & \mathrm{GSp}_8(F) \\
& (a, b, c) & \mapsto & a \otimes b \otimes c,
\end{array} \tag{7.6}$$

where we identify F^8 with $F^2 \otimes F^2 \otimes F^2$. We equip F^8 with the symplectic form ψ given by $\psi_1 \otimes \psi_2 \otimes \psi_3$, where ψ_i is the standard symplectic form on the i-th factor F^2: the fact that every element of $\mathrm{GL}_2(F)$ preserves ψ_i (up to a scalar) implies that $\rho(a, b, c)$ preserves ψ (up to a scalar), so the image of ρ is indeed contained in $\mathrm{GSp}_8(F)$.

Definition 7.3.1. We let M_F be the F-Zariski closure of the image of this representation (with its obvious structure as an algebraic group over F).

Remark 7.3.2. For all $\lambda \in F^\times$ the matrix $\lambda \cdot \mathrm{Id}$ belongs to the image of ρ. In particular, if F is an infinite field the group M_F contains the (algebraic) group of homotheties.

Proposition 7.3.3. *For every prime ℓ we have $M_{\mathbb{Q}_\ell} \cong M_{\mathbb{Q}} \times_{\mathbb{Q}} \mathbb{Q}_\ell$.*

Proof. The inclusion $M_{\mathbb{Q}} \otimes \mathbb{Q}_\ell \subseteq M_{\mathbb{Q}_\ell}$ is obvious (since with $M_{\mathbb{Q}_\ell}$ we are taking the Zariski closure of a larger set). On the other hand it is not hard to see that both groups become isomorphic to $\mathbb{G}_m \cdot \mathrm{SL}_2^3$ over an algebraic closure, so they have the same dimension and therefore must coincide. □

Remark 7.3.4. Consider the $\mathbb{Z}_\ell$-Zariski closure of $M_{\mathbb{Q}_\ell}$ in $\mathrm{GSp}_{8,\mathbb{Z}_\ell}$, call it $\mathcal{M}_{\mathbb{Z}_\ell}$. In view of the proposition, $\mathcal{M}_{\mathbb{Z}_\ell}$ coincides with the $\mathbb{Z}_\ell$-Zariski closure of $M_{\mathbb{Q}} \times_{\mathbb{Q}} \mathbb{Q}_\ell$ in $\mathrm{GSp}_{8,\mathbb{Z}_\ell}$, and the latter is smooth over $\mathbb{Z}_\ell$ for almost all ℓ because $M_{\mathbb{Q}}$ extends to a smooth scheme over an open subscheme of $\operatorname{Spec}\mathbb{Z}$. It follows that that $\mathcal{M}_{\mathbb{Z}_\ell}$ is smooth over $\mathbb{Z}_\ell$ for almost all ℓ.

We think the algebraic group M_F as sitting inside $\mathbb{A}_F^{64} = M_8(F)$. It is not hard to find polynomials that vanish identically on the image of ρ: indeed, if we let $\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ be any element in $\operatorname{Im}\rho$ (where every B_{ij} is a 4×4 matrix), the construction of the tensor product implies that the four matrices B_{ij} are pairwise linearly dependent, a condition which is purely algebraic (being given by the vanishing of sufficiently many determinants); in particular, the same property is valid for *any* matrix in $M_F(\overline{F})$. Likewise, if we write $B_{ij} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$, where each C_{kl} is a 2×2 matrix, we must again have pairwise linear dependence of the C_{kl}'s, and this (being an algebraic condition) is again true for any point in $M_F(\overline{F})$. Let now e_1, e_2 be the standard basis of F^2 and write $e_{ijk} = e_i \otimes e_j \otimes e_k$ (with $i,j,k \in \{1,2\}$) for the corresponding basis of F^8. We order these basis vectors as $e_{111}, e_{112}, e_{121}, e_{122}, e_{211}, e_{212}, e_{221}, e_{222}$. The form ψ on F^8 is then represented by the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and it is immediate to check that $e_{111}, e_{122}, e_{212}, e_{221}$ span a Lagrangian subspace.

Definition 7.3.5. Let F be any field. We let H be the subspace of $F^8 \cong \left(F^2\right)^{\otimes 3}$ generated by e_{111}, e_{122}, e_{212}, and e_{221}.

We now determine the stabilizer S of H in $M_F\left(\overline{F}\right)$. In matrix terms, an element s of S can be written as

$$s = \begin{pmatrix} 1 & \square & \square & 0 & \square & 0 & 0 & \square \\ 0 & \square & \square & 0 & \square & 0 & 0 & \square \\ 0 & \square & \square & 0 & \square & 0 & 0 & \square \\ 0 & \square & \square & 1 & \square & 0 & 0 & \square \\ 0 & \square & \square & 0 & \square & 0 & 0 & \square \\ 0 & \square & \square & 0 & \square & 1 & 0 & \square \\ 0 & \square & \square & 0 & \square & 0 & 1 & \square \\ 0 & \square & \square & 0 & \square & 0 & 0 & \square \end{pmatrix},$$

where each entry $\square$ is a priori any element of $\overline{F}$. We now use the fact that $S \subseteq M_F(\overline{F})$ to show that S is finite. Write as before B_{11} (resp. B_{12}, B_{21}, B_{22}) for the top-left (resp. top-right, bottom-left and bottom-right) block of s of size 4×4. Since B_{22} is nonzero, linear dependence of B_{22} and B_{12} can be expressed as $B_{12} = \alpha B_{22}$ for a certain $\alpha \in \overline{F}$; however, since B_{22} has some nonzero diagonal coefficients while the corresponding diagonal entries of B_{12} vanish, we must have $\alpha = 0$ and $B_{12} = 0$. The same argument, applied to B_{21} and B_{11}, shows that $B_{21} = 0$. On the other hand, the blocks B_{11} and B_{22} are both nonzero, so there exists a nonzero $\lambda \in \overline{F}^\times$ such that $B_{22} = \lambda B_{11}$: this leads immediately to

$$s = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/\lambda & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/\lambda & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda \end{pmatrix}.$$

We now use the second part of our previous remark, namely the fact that the 2×2 blocks of B_{11} are linearly dependent as well. Comparing the top-left and bottom-right blocks of B_{11} gives the additional condition $\lambda^2 = 1$, that is, $\lambda = \pm 1$: thus the stabilizer in $M_F(\overline{F})$ of our Lagrangian subspace H consists of exactly two elements, namely the identity and the operator $\mathrm{diag}(1, -1, -1, 1, -1, 1, 1, -1)$ (at least if $\mathrm{char}\, F \neq 2$: otherwise we have $-1 = 1$ and the two coincide). This stabilizer is also clearly finite as an algebraic group, since it has only finitely many points over $\overline{F}$.

Notice that this argument actually shows a little more. Let $\mathcal{M}_{\mathbb{Z}_\ell}$ be the $\mathbb{Z}_\ell$-Zariski closure of $M_{\mathbb{Q}_\ell}$ in $\mathrm{GSp}_{8,\mathbb{Z}_\ell}$, and suppose that $\mathcal{M}_{\mathbb{Z}_\ell}$ is smooth over $\mathbb{Z}_\ell$. Let furthermore $\mathcal{H}$ be the Lagrangian subspace of $\mathbb{F}_\ell^8 \cong \mathbb{F}_\ell^2 \otimes \mathbb{F}_\ell^2 \otimes \mathbb{F}_\ell^2$ given in definition 7.3.5 (for the field $\mathbb{F}_\ell$): then the stabilizer of $\mathcal{H}$ in $\mathcal{M}_{\mathbb{Z}_\ell}(\overline{\mathbb{F}_\ell})$ has order at most 2. Indeed, all we have used to show that $|M_F(\overline{F})| \leq 2$ is the linear dependence of certain blocks in the matrix representation of its elements and the fact that the equation $\lambda^2 = 1$ admits at most 2 solutions in $\overline{F}$: both properties are also true for the points of $\mathcal{M}_{\mathbb{Z}_\ell}$ with values in any integral $\mathbb{Z}_\ell$-algebra (in particular, $\overline{\mathbb{F}_\ell}$). We record this fact in the following

Proposition 7.3.6. *Let ℓ be a prime, $\mathcal{M}_{\mathbb{Z}_\ell}$ be the $\mathbb{Z}_\ell$-Zariski closure of $M_{\mathbb{Q}_\ell}$ in $\mathrm{GSp}_{8,\mathbb{Z}_\ell}$, and $\mathcal{H}$ be the subspace H of definition 7.3.5 for the field $\mathbb{F}_\ell$. The stabilizer of $\mathcal{H}$ in $\mathcal{M}_{\mathbb{Z}_\ell}(\overline{\mathbb{F}_\ell})$ consists of at most 2 elements.*

7.3.1 Mumford's examples

We now recall the construction given by Mumford in [85]. Suppose we are given the data of a totally real cubic number field F and of a central simple division algebra D over F satisfying:

1. $\mathrm{Cor}_{F/\mathbb{Q}}(D) = M_8(\mathbb{Q})$;
2. $D \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{H} \oplus \mathbb{H} \oplus M_2(\mathbb{R})$.

Being a division algebra, D is equipped with a natural involution $x \mapsto \overline{x}$; let G be the $\mathbb{Q}$-algebraic group whose $\mathbb{Q}$-points are given by $\left\{x \in D^* \;\middle|\; x\overline{x} = 1\right\}$. Mumford constructed in [85] an abelian variety of dimension 4 with trivial endomorphism ring and Hodge group equal to G (in fact, he constructed a Shimura curve parametrizing abelian fourfolds whose Hodge group is contained in G, and showed that every sufficiently generic fiber has exactly G as its Hodge group). By specialization, there exists a principally polarized abelian fourfold A defined over a number field L and such that $\mathrm{Hg}(A) \cong G$; since $\mathrm{Hg}(A)$ is as small as it is possible for an abelian fourfold with no additional endomorphisms, the Mumford-Tate conjecture is known to hold for A (cf. [80]). By theorem 7.2.3 there is a finite extension K of L such that, if we denote G_ℓ the image of the mod-ℓ representation $\mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}\, A[\ell]$, then we have $G_\ell \subseteq \mathrm{MT}(A)(\mathbb{F}_\ell)$ for all primes ℓ. On the other hand, the equality $\mathrm{Cor}_{F/\mathbb{Q}}(D) = M_8(\mathbb{Q})$ implies the existence of a ("norm") map $N : D^* \to \mathrm{GL}_8(\mathbb{Q})$, and Mumford's construction is such that the action of $G(\mathbb{Q}) = D^*$ on $V := H_1(A(\mathbb{C}), \mathbb{Q}) \cong \mathbb{Q}^8$ is given exactly by N. Furthermore, it is also known that N is a $\mathbb{Q}$-form of the $\mathbb{R}$-representation

$$G(\mathbb{R}) \cong \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SU}_2(\mathbb{R})^2 \to \mathrm{Sp}_8(\mathbb{R})$$

coming from the tensor product of the standard representation of $\mathrm{SL}_2(\mathbb{R})$ by the unique four-dimensional faithful orthogonal representation $\mathrm{SU}_2(\mathbb{R})^2 \to \mathrm{SO}_4(\mathbb{R})$. In particular, by extension of scalars to $\mathbb{C}$ we see that the action of $G(\mathbb{C}) \cong \mathrm{SL}_2(\mathbb{C})^3$ on $V_{\mathbb{C}}$ is given by the representation ρ of the previous paragraph (restricted to $\mathrm{SL}_2(\mathbb{C})^3$).

Lemma 7.3.7. *Let ℓ be a prime such that $G \times_{\mathbb{Q}} \mathbb{Q}_\ell$ is split. Then (up to choosing a suitable identification $T_\ell(A) \otimes \mathbb{Q}_\ell \cong \mathbb{Q}_\ell^8$) we have $\mathrm{MT}(A) \times_{\mathbb{Z}} \mathbb{Q}_\ell = M \times_{\mathbb{Q}} \mathbb{Q}_\ell$, where $M = M_{\mathbb{Q}}$ is the algebraic group of definition 7.3.1 for the field $\mathbb{Q}$.*

Proof. The morphism $G \to \mathrm{Sp}_{8,\mathbb{Q}}$ is given by the norm map, and if $G \times_{\mathbb{Q}} \mathbb{Q}_\ell$ is split (hence isomorphic to $\mathrm{SL}_{2,\mathbb{Q}_\ell}^3$) the norm map is exactly

$$\begin{array}{rccc} \rho : & \mathrm{SL}_2(\mathbb{Q}_\ell)^3 & \to & \mathrm{Sp}_8(\mathbb{Q}_\ell) \\ & (a,b,c) & \mapsto & a \otimes b \otimes c; \end{array}$$

it follows that $M(\mathbb{Q}_\ell)$ contains $\rho\left(G(\mathbb{Q}_\ell)\right) = \mathrm{Hg}(A)(\mathbb{Q}_\ell)$ and $M \times_{\mathbb{Q}} \mathbb{Q}_\ell$ contains $\mathrm{Hg}(A) \times_{\mathbb{Q}} \mathbb{Q}_\ell$ (as algebraic groups). On the other hand, $\mathrm{MT}(A)$ is the almost-direct product of $\mathrm{Hg}(A)$ by the homotheties torus $\mathbb{G}_m$, and by remark 7.3.2 we know that M also contains $\mathbb{G}_m$. This proves that we have $\mathrm{MT}(A) \times \mathbb{Q}_\ell \subseteq M \times \mathbb{Q}_\ell$, and since the two groups have the same dimension the inclusion must be an equality. □

Extend now M and G to group schemes over $\mathbb{Z}$ by taking their $\mathbb{Z}$-Zariski closure in their respective ambient spaces; there is an open subscheme $\operatorname{Spec}\mathbb{Z}\left[\frac{1}{S}\right]$ of $\operatorname{Spec}\mathbb{Z}$ over which $M, \mathrm{MT}(A)$ and G are all smooth. Consider the family $\mathcal{F}$ of primes ℓ unramified in K, such that G splits over $\mathbb{Q}_\ell$, and which do not divide S. We claim that $\mathcal{F}$ is infinite. Indeed, for G to be split over $\mathbb{Q}_\ell$ it is enough that the root datum of G be unramified at ℓ and that the Frobenius at ℓ act trivially on it, which – by Chebotarev's theorem – is the case for a positive-density set of primes (the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the root datum of G factors through a finite quotient): it is then clear that $\mathcal{F}$ is infinite, because only finitely many primes divide S or the discriminant of K. Pick now any ℓ in $\mathcal{F}$ and let $\mathcal{M} = M \times_{\mathbb{Z}} \mathbb{Z}_\ell$. The definition of $\mathcal{F}$ implies that $\mathcal{M}$ is a smooth $\mathbb{Z}_\ell$-model of $M \times_{\mathbb{Z}} \mathbb{Q}_\ell = M_{\mathbb{Q}_\ell}$, and by lemma 7.3.7 we have $\mathrm{MT}(A) \times_{\mathbb{Z}} \mathbb{Z}_\ell = \mathcal{M}$, because both groups can be obtained as the $\mathbb{Z}_\ell$-Zariski closure of the same generic fiber. In particular, we see that G_ℓ is contained in $\mathcal{M}(\mathbb{F}_\ell) = \mathrm{MT}(A)(\mathbb{F}_\ell)$. Take now $\mathcal{H} \subseteq A[\ell]$ to be the Lagrangian subspace of definition 7.3.5 (for the field $\mathbb{F}_\ell$). The field $K(\mathcal{H})$ is clearly contained in $K(A[\ell])$, so in order to describe $K(\mathcal{H})$ it suffices to describe $\mathrm{Gal}\left(K(A[\ell])/K(\mathcal{H})\right)$, that is, the stabilizer of $\mathcal{H}$ in G_ℓ; as G_ℓ is contained in $\mathcal{M}(\mathbb{F}_\ell)$, this stabilizer is certainly contained in the stabilizer of $\mathcal{H}$ in $\mathcal{M}(\mathbb{F}_\ell)$, which in turn consists of at most two elements by proposition 7.3.6. We have thus proved that the index $[K(A[\ell]) : K(\mathcal{H})]$ is at most 2, and since $K(\mu_\ell)$ is contained in $K(A[\ell])$ by the properties of the Weil pairing (recall that A is principally polarized) we have

$$[K(\mathcal{H}) \cap K(\mu_{\ell^\infty}) : K] \geq \frac{1}{2}\left[K(A[\ell]) \cap K(\mu_{\ell^\infty}) : K\right] \geq \frac{1}{2}\left[K(\mu_\ell) : K\right] = \frac{\ell - 1}{2},$$

where the last equality follows from the fact that ℓ is unramified in K. We then see that property $(\mu)_s$ does not hold for Mumford's example: indeed, $\mathcal{H}$ is Lagrangian, so $m_1(\mathcal{H})$ is 0; but if property $(\mu)_s$ held for A/K, then (for some C) the inequality

$$\frac{\ell - 1}{2} \leq [K(\mathcal{H}) \cap K(\mu_{\ell^\infty}) : K] \leq C\left[K(\mu_{\ell^{m_1(\mathcal{H})}}) : K\right] = C$$

would be satisfied by all the primes in our infinite family $\mathcal{F}$, and this is clearly absurd. This establishes theorem 7.1.5.

Chapter 8

Pink-type results for general subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$

8.1 Motivation and statement of the result

The ultimate goal of this work is the study of the images of certain Galois representations with values in $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$, such as those afforded by the Tate modules of elliptic curves, or some representations arising from modular forms. It would therefore be useful to have a manageable way to describe these images; however, it turns out that it is beneficial, and in a sense simpler, to consider arbitrary subgroups G of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$ without making any reference to their origin, and in the present work Galois representations will play virtually no role. In most applications to the study of Galois representations, the main object of interest is actually the intersection $G \cap \mathrm{SL}_2(\mathbb{Z}_\ell)^n$, and furthermore it is an easy matter to pass from results on subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$ to results on subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$, so we shall actually mostly work with subgroups G of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$. Any such G is the extension of a 'finite' part, the image $G(\ell)$ of the reduction $G \to \mathrm{SL}_2(\mathbb{F}_\ell)^n$, by a 'Lie' part, the kernel of this reduction.

When G is closed and $G(\ell)$ is trivial (or more generally when G is pro-ℓ), and ℓ is odd, a construction due to Pink [97] gives a very concrete and handy description of G in terms of a certain $\mathbb{Z}_\ell$-Lie algebra $L(G)$ (together with some additional data which is not very important to our present discussion). Furthermore, if G is the image of a representation of $\mathrm{Gal}(\overline{K}/K)$ (K a number field), the condition that $G(\ell)$ be trivial can always be met by replacing K by a finite extension, so that Pink's theorem applies. Note however that the degree of this extension depends on ℓ: while this is often perfectly fine when considering a single Galois representation, it may become a major drawback when dealing with infinite families G_ℓ indexed by the rational primes (as it is the case, for example, with the action of $\mathrm{Gal}(\overline{K}/K)$ on the various Tate modules of an abelian variety). Furthermore, Pink's theorem does not apply to $\ell = 2$, which might again be quite a hindrance when trying to study the whole system G_ℓ at once.

While we cannot hope to give a complete description of G in terms of Pink's Lie algebras when G is not pro-ℓ, we could try and settle for less, namely a result of the form 'when $L(G)$ contains a large neighbourhood of the identity (given explicitly), we can explicitly find a neighbourhood of the identity of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$ that is included in G'. Note that when dealing with Galois representations

we are often interested in 'large image' results, for which this weaker form of Pink's theorem would still be adequate. Unfortunately, even this is not possible (cf. for example §1.4.5), and the best we can hope for is for such a statement to hold not quite for G, but for a subgroup H of G such that the index $[G : H]$ is bounded *by a function of* n *alone*.
In order to give a concrete statement we shall need some preliminary definitions:

Definition 8.1.1. For a prime ℓ and a positive integer s we let $\mathcal{B}_\ell(s)$ be the open subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ given by

$$\left\{x \in \mathrm{SL}_2(\mathbb{Z}_\ell) \;\middle|\; x \equiv \mathrm{Id} \pmod{\ell^s}\right\}.$$

We also set $\mathcal{B}_\ell(0) = \mathrm{SL}_2(\mathbb{Z}_\ell)$, and for non-negative integers $k_1, \ldots, k_n$ we denote by $\mathcal{B}_\ell(k_1, \ldots, k_n)$ the open subgroup $\prod_{j=1}^n \mathcal{B}_\ell(k_j)$ of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$.

Definition 8.1.2. (cf. [97]) Let ℓ be a prime, n a positive integer and G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)^n$. If $\ell = 2$, assume further that the reduction modulo 4 of G is trivial. Writing elements of G as n-tuples $(g_1, \ldots, g_n)$ of elements of $\mathrm{GL}_2(\mathbb{Z}_\ell)$, we define a map Θ_n by the formula

$$\begin{array}{rccc} \Theta_n : & G & \to & \bigoplus_{i=1}^n \mathfrak{sl}_2(\mathbb{Z}_\ell) \\ & (g_1, \ldots, g_n) & \mapsto & \left(g_1 - \frac{1}{2}\operatorname{tr}(g_1), \ldots, g_n - \frac{1}{2}\operatorname{tr}(g_n)\right), \end{array}$$

and we let $L(G) \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)^n$ be the $\mathbb{Z}_\ell$-span of $\Theta_n(G)$. We call $L(G)$ the Lie algebra of G.

Theorem 8.1.3. *Let ℓ be an odd prime, n be an integer, and G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)^n$. There exists a closed subgroup H of G, of index at most $24^n 48^{n(n-1)}$, with the following property: if $L(H)$ contains $\bigoplus_{i=1}^n \ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell)$ for a certain integer $k > 0$, then H contains $B_\ell(p, \ldots, p)$ for $p = 80(\max\{n, 2\} - 1)k$.*
Similarly, let n be a positive integer and G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)^n$. There exists a closed subgroup H of G that satisfies $[G : H] \mid 96^n$, is trivial modulo 4 (so that $L(H)$ is defined), and has the following property: if $L(H)$ contains $\bigoplus_{i=1}^n 2^k \mathfrak{sl}_2(\mathbb{Z}_2)$ for a certain integer $k > 0$, then H contains $B_2(p, \ldots, p)$ for $p = 645(\max\{n, 2\} - 1)k$.

While it is certainly true that both this theorem and its proof are quite technical, it should be remarked that this statement does enable us to show exactly the kind of 'large image' results we alluded to: the case $n = 1$ has been used in chapter 1 to show an explicit open image theorem for elliptic curves (without complex multiplication), and in chapter 2 we apply the case $n = 2$ to extend this result to arbitrary products of non-CM elliptic curves.
A few more words on the proof of theorem 8.1.3: as it will be clear from section 8.5, the crucial cases are $n = 1$ and $n = 2$. While the former has essentially been proven in chapter 1, the latter forms the core of the present chapter, and we shall actually prove it in a slightly more precise form than strictly necessary to establish theorem 8.1.3. This will be done in sections 8.3 and 8.4 below, where we also give analogous statements for $\mathrm{GL}_2(\mathbb{Z}_\ell)^2$.

Notation. We shall make constant use of the following notations:

- $G(\ell^n)$, where G is a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)^k$, will denote the reduction of G modulo ℓ^n, that is to say its image in $\mathrm{GL}_2(\mathbb{Z}/\ell^n\mathbb{Z})^k$;

- $N(G)$ will denote the largest normal pro-ℓ subgroup of G, where G is a subgroup of either $\mathrm{GL}_2(\mathbb{Z}_\ell)$ or of $\mathrm{GL}_2(\mathbb{F}_\ell)$;
- G' will denote the topological closure of the commutator subgroup;
- if x is an element of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ (resp. of $\mathbb{Z}_\ell$), we shall write $[x]$ for its image in $\mathrm{GL}_2(\mathbb{F}_\ell)$ (resp. in $\mathbb{F}_\ell$).

8.2 Preliminary lemmas

In this section we set up the necessary notation and prove a few preliminary lemmas.

Throughout the whole chapter, the prime 2 plays a rather special role, and special care is needed to treat it. In order to give uniform statements that hold for every prime, we put $v = 0$ or 1 according to whether the prime ℓ we are working with is odd or equals 2, that is we set

$$v = v_\ell(2) = \begin{cases} 0, \text{ if } \ell \text{ is odd} \\ 1, \text{ otherwise.} \end{cases}$$

Lemma 8.2.1. *Let ℓ be a prime number, t a non-negative integer, and $W \subseteq \mathfrak{sl}_2(\mathbb{Z}_\ell)$ a Lie subalgebra that does not reduce to zero modulo ℓ^{t+1} and that is stable under conjugation by $\mathcal{B}_\ell(s)$, where $s \geq 0$ is at least 2 if $\ell = 2$ and at least 1 if $\ell = 3$ or 5 (no conditions are necessary if $\ell \geq 7$). The open set $\ell^{t+4s+4v}\mathfrak{sl}_2(\mathbb{Z}_\ell)$ is contained in W.*

Proof. Write an element w of W that does not vanish modulo ℓ^{t+1} as $\mu_x x + \mu_y y + \mu_h h$, where

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and $\min\{v_\ell(\mu_x), v_\ell(\mu_y), v_\ell(\mu_h)\} \leq t$.

Let $\mathcal{C}_l$ (resp. $\mathcal{C}_d, \mathcal{C}_r$) be the linear operator of W into itself given by conjugation by $\begin{pmatrix} 1 & 0 \\ \ell^s & 1 \end{pmatrix}$ (resp. $\begin{pmatrix} 1+\ell^s & 0 \\ 0 & \frac{1}{1+\ell^s} \end{pmatrix}, \begin{pmatrix} 1 & \ell^s \\ 0 & 1 \end{pmatrix}$) and $\mathcal{D}_\bullet = \mathcal{C}_\bullet - \mathrm{Id}$, where $\bullet$ is one among l, d, r. Concretely,

$$\mathcal{D}_l(w) = \begin{pmatrix} 1 & 0 \\ \ell^s & 1 \end{pmatrix}^{-1} w \begin{pmatrix} 1 & 0 \\ \ell^s & 1 \end{pmatrix} - w.$$

Also set $\alpha := 1 + \ell^s$. We have

$$\alpha^4(\mathcal{C}_d - \alpha^2) \circ \mathcal{D}_d w = (\alpha^4 - 1)(\alpha^2 - 1)\mu_x x \in W,$$

where $v_\ell((\alpha^4 - 1)(\alpha^2 - 1)) = 2s + 3v$ by our assumptions on s. Similarly, we also have

$$(\alpha^4 - 1)(\alpha^2 - 1)\mu_y y \in W$$

and by difference also $(\alpha^4 - 1)(\alpha^2 - 1)\mu_h h \in W$. Up to symmetry, we can therefore assume that W contains either $M_1 = \begin{pmatrix} \ell^{t+2s+3v} & 0 \\ 0 & -\ell^{t+2s+3v} \end{pmatrix}$ or $M_2 = \begin{pmatrix} 0 & \ell^{t+2s+3v} \\ 0 & 0 \end{pmatrix}$. To finish the proof we use the following immediate identities:

- $\mathcal{D}_r(M_1) = \begin{pmatrix} 0 & 2\ell^{t+3s+3v} \\ 0 & 0 \end{pmatrix}, \mathcal{D}_l(M_1) = \begin{pmatrix} 0 & 0 \\ -2\ell^{t+3s+3v} & 0 \end{pmatrix}$, so that in this case W contains $\ell^{t+3s+4v}\mathfrak{sl}_2(\mathbb{Z}_\ell)$
- $\mathcal{D}_l \circ \mathcal{D}_l(M_2) = \begin{pmatrix} 0 & 0 \\ -2\ell^{t+4s+3v} & 0 \end{pmatrix}, \mathcal{D}_l \circ \mathcal{D}_l(M_2) - 2\mathcal{D}_l(M_2) = \begin{pmatrix} -2\ell^{t+3s+3v} & 0 \\ 0 & 2\ell^{t+3s+3v} \end{pmatrix}$, so that in this case W contains $\ell^{t+4s+4v}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

□

Lemma 8.2.2. *Let ℓ be a prime number, $n \geq 1, m \geq 1$, $g \in \mathrm{End}(\mathbb{Z}_\ell^m)$ and $p_g(t)$ be the characteristic polynomial of g. Let furthermore $\lambda \in \mathbb{Z}_\ell, w \in \mathbb{Z}_\ell^m$ be such that $gw \equiv \lambda w \pmod{\ell^n}$. Suppose that at least one of the coordinates of w has ℓ-adic valuation at most α: then $p_g(\lambda) \equiv 0 \pmod{\ell^{n-\alpha}}$.*

Proof. Denote $(g - \lambda \mathrm{Id})^*$ the adjugate matrix of $(g - \lambda \mathrm{Id})$, that is the unique operator such that $(g - \lambda \mathrm{Id})^*(g - \lambda \mathrm{Id}) = \det(g - \lambda \mathrm{Id}) \cdot \mathrm{Id}$. Multiplying $(g - \lambda \mathrm{Id})w \equiv 0 \pmod{\ell^n}$ on the left by $(g - \lambda \mathrm{Id})^*$ we obtain $\det(g - \lambda \mathrm{Id}) \cdot \mathrm{Id}\, w \equiv 0 \pmod{\ell^n}$, and by considering the coordinate of w of smallest valuation we obtain $p_g(\lambda) = \det(g - \lambda \mathrm{Id}) \equiv 0 \pmod{\ell^{n-\alpha}}$ as claimed. □

Lemma 8.2.3. *Let s_1, s_2 be non-negative integers (with $s_1, s_2 \geq 2$ if $\ell = 2$ and $s_1, s_2 \geq 1$ if $\ell = 3$). The commutator group $[\mathcal{B}_\ell(s_1), \mathcal{B}_\ell(s_2)]$ contains $B_\ell(s_1+s_2+v)$, and the iterated commutator $\underbrace{[\cdots[}_{(n-1) \text{ times}} \mathcal{B}_\ell(s_1), \mathcal{B}_\ell(s_2)\,], \mathcal{B}_\ell(s_3)], \cdots, \mathcal{B}_\ell(s_n)]$ contains $B_\ell(s_1 + \cdots + s_n + (n-1)v)$.*

Proof. This is an easy verification. □

The quantitative result we will need is the following:

Lemma 8.2.4. *Let n be a positive integer, G a closed subgroup of $\prod_{i=1}^n \mathrm{SL}_2(\mathbb{Z}_\ell)$, and π_i the projection from G on the i-th factor. Suppose that, for every $i \neq j$, the group $(\pi_i \times \pi_j)(G)$ contains $\mathcal{B}_\ell(s_{ij}, s_{ij})$ for a certain non-negative integer s_{ij} (with $s_{ij} \geq 2$ if $\ell = 2$ and $s_{ij} \geq 1$ if $\ell = 3$): then G contains $\prod_{i=1}^n \mathcal{B}_\ell\left(\sum_{j \neq i} s_{ij} + (n-1)v\right)$.*

Proof. Clearly by the symmetry of the problem it is enough to show that G contains

$$\{\mathrm{Id}\} \times \cdots \times \{\mathrm{Id}\} \times \mathcal{B}_\ell\left(\sum_{j \neq n} s_{nj} + (n-1)v\right).$$

By lemma 8.2.3, for any g in $\mathcal{B}_\ell\left(\sum_{j \neq n} s_{nj} + (n-1)v\right)$ there exist elements y_i in $\mathcal{B}_\ell(s_{ni})$ (for i between 1 and $n-1$) such that g can be written as $[\cdots[[y_1, y_2], y_3], \cdots, y_{n-1}]$. By hypothesis we can find $x_1, \ldots, x_{n-1} \in G$ such that $\pi_i(x_i) = \mathrm{Id}$ and $\pi_n(x_i) = y_i$ for all i between 1 and $n-1$. Consider now the iterated commutator

$$\tilde{g} = [\cdots[[x_1, x_2], x_3], \cdots, x_{n-1}]:$$

this is a product of elements of G, and therefore it is itself an element of G. For $i \leq n-1$, the i-th component of $\tilde{g}$ is trivial, since

$$\pi_i(\tilde{g}) = [\cdots[\cdots[[\pi_i(x_1), \pi_i(x_2)], \pi_i(x_3)], \cdots, \underbrace{\pi_i(x_i)}_{\mathrm{Id}}], \cdots, \pi_i(x_{n-1})] = \mathrm{Id}\,.$$

On the other hand, our choice of $y_1, \ldots, y_{n-1}$ ensures that $\pi_n(\tilde{g}) = [\cdots[[y_1, y_2], y_3], \cdots, y_{n-1}] = g$. We have thus shown that $(1, 1, \ldots, 1, g) = \tilde{g}$ is an element of G for any choice of g in

$$\mathcal{B}_\ell\left(\sum_{j\neq n} s_{ij} + (n-1)v\right),$$

and repeating the argument for the other projections gives the required result. □

8.3 Odd ℓ, $n = 2$

In this section we establish the case $n = 2$ of the main theorem when ℓ is an odd prime. We shall actually prove the following variant concerning subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)^2$:

Theorem 8.3.1. *Let $\ell > 2$ be a prime number and G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell) \times \mathrm{GL}_2(\mathbb{Z}_\ell)$. Let G_1, G_2 be the two projections of G on the two factors $\mathrm{GL}_2(\mathbb{Z}_\ell)$, and let n_1, n_2 be positive integers such that G_i contains $\mathcal{B}_\ell(n_i)$ for $i = 1, 2$. Suppose furthermore that for every $(g_1, g_2) \in G$ we have $\det(g_1) = \det(g_2)$. At least one of the following holds:*

- *G contains $\mathcal{B}_\ell(20\max\{n_1, n_2\}, 20\max\{n_1, n_2\})$*
- *there exists a subgroup T of G, of index dividing $2 \cdot 48^2$, with the following properties:*
  - *if $L(T)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$ for a certain integer k, then T contains $\mathcal{B}_\ell(p, p)$, where*

    $$p = 2k + \max\{2k + 4, 8n_1, 8n_2\}.$$

    We call this property $(*)$.
  - *for any (t_1, t_2) in T, if both $[t_1]$ and $[t_2]$ are multiples of the identity, then they are equal.*

The corresponding statement for subgroups of $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$ is as follows:

Theorem 8.3.2. *Let $\ell > 2$ be a prime number and G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell) \times \mathrm{SL}_2(\mathbb{Z}_\ell)$ and n_1, n_2 be positive integers such that G_i contains $\mathcal{B}_\ell(n_i)$ for $i = 1, 2$. At least one of the following holds:*

- *G' contains $\mathcal{B}_\ell(20\max n_1, n_2, 20\max n_1, n_2)$*
- *there exists a subgroup T of G, of index dividing 48^2, with the following properties:*
  - *if $L(T)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$ for a certain integer k, then T' contains $\mathcal{B}_\ell(p, p)$, where*

    $$p = 2k + \max\{2k + 4, 8n_1, 8n_2\}.$$

    We call this property $(**)$.
  - *an element (t_1, t_2) of T satisfies $[t_1] = [\mathrm{Id}]$ if and only if it satisfies $[t_2] = [\mathrm{Id}]$.*

We will start by showing that theorem 8.3.2 implies theorem 8.3.1, and then proceed to prove the former. One of the key ingredients of the proof is the following theorem, which in turn is an immediate consequence of Pink's results from [97] (see also the proof of theorem 1.4.2).

Theorem 8.3.3. *Let $\ell \neq 2$ be a prime number and k be an integer.*
Suppose G is a closed pro-ℓ subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$, and suppose furthermore that $L(G)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$: then G' contains $\mathcal{B}_\ell(2k)$. Similarly, if G is a closed pro-ℓ subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$ and $L(G)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$, then G' contains $\mathcal{B}_\ell(2k, 2k)$.

Proof. (of theorem 8.3.1) Write $\det^*$ for the map $G \xrightarrow{\det} \mathbb{Z}_\ell^2 \xrightarrow{\pi_1} \mathbb{Z}_\ell$ given by the composition of the usual determinant with the projection on the first coordinate of $\mathbb{Z}_\ell^2$. Note that by assumption an element (g_1, g_2) of G satisfies $\det(g_1, g_2) = (1, 1)$ if and only if it satisfies $\det^*(g_1, g_2) = 1$.
Assume first that $\ell \leq 5$. Denote by $\tilde{T}$ the inverse image in G of an ℓ-Sylow of $G(\ell)$, and set

$$T := \ker\left(\tilde{T} \xrightarrow{\det^*} \mathbb{Z}_\ell^\times \to \frac{\mathbb{Z}_\ell^\times}{\mathbb{Z}_\ell^{\times 2}}\right).$$

By the assumption that every element (g_1, g_2) of G satisfies $\det g_1 = \det g_2$, we see that the index of T in G divides $2 \cdot \frac{1}{\ell-1}\left(\frac{|\mathrm{GL}_2(\mathbb{F}_\ell)|}{\ell}\right)^2 \mid 2 \cdot 48^2$. As in lemma 1.3.17, one sees that the groups T and $T^1 := \left(T \cdot (\mathbb{Z}_\ell^\times \cdot (\mathrm{Id}, \mathrm{Id}))\right) \cap \mathrm{SL}_2(\mathbb{Z}_\ell)^2$ have the same derived subgroup and the same Lie algebra, and moreover T^1 is a pro-ℓ subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$. Furthermore, every element (t_1, t_2) of T reduces to $([\mathrm{Id}], [\mathrm{Id}])$ modulo ℓ. Now if $L(T)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$, then the same is true for $L\left(T^1\right)$, hence $(T^1)' = T'$ contains $\mathcal{B}_\ell(2k, 2k)$ by theorem 8.3.3.
Next consider the case $\ell > 5$. Let U_1 be the subgroup of G, of index at most 2, given by $\ker\left(G \xrightarrow{\det^*} \mathbb{Z}_\ell^\times \to \mathbb{Z}_\ell^\times/\mathbb{Z}_\ell^{\times 2}\right)$. Let $U_2 = \left(U_1 \cdot \left(\mathbb{Z}_\ell^\times \cdot (\mathrm{Id}, \mathrm{Id})\right)\right) \cap \mathrm{SL}_2(\mathbb{Z}_\ell)^2$, and notice that U_1 and U_2 have the same derived subgroup (lemma 1.3.17). We can assume that $U_2' = U_1'$ does not contain $\mathcal{B}_\ell(4n_1 + 16n_2, 8n_2)$, for otherwise we are done. Apply theorem 8.3.2 to U_2 to find a subgroup T_2 of U_2 (of index at most 48^2) that has property $(**)$. Notice that we have a well-defined morphism

$$U_1 \xrightarrow{\psi} U_2/(\pm(\mathrm{Id}, \mathrm{Id}))$$

given by $g \mapsto [g/\sqrt{\det^*(g)}]$, where $\sqrt{\det^* g}$ exists in $\mathbb{Z}_\ell^\times$ by construction of U_1. Let $\overline{T_2}$ be the image of T_2 in the quotient $U_2/(\pm(\mathrm{Id}, \mathrm{Id}))$. Notice that ψ is surjective by definition of U_2, so if we define T to be the inverse image of $\overline{T_2}$ through ψ, then the index $[U_1 : T]$ divides 48^2. As the prime ℓ is larger than 5, it does not divide $[G : T]$, so (given that $\mathcal{B}_\ell(n_1), \mathcal{B}_\ell(n_2)$ are ℓ-groups) the two projections of T on the two factors $\mathrm{SL}_2(\mathbb{Z}_\ell)$ contain $\mathcal{B}_\ell(n_1), \mathcal{B}_\ell(n_2)$ respectively. Furthermore, the Lie algebra of T and that of T_2 agree, as do their derived subgroups: for every $t_2 \in T_2$ we can find a $t \in T$ such that $t/\sqrt{\det^* t} = \pm t_2$, so that $\Theta_1(t)$ and $\Theta_1(t_2)$ differ by an element of $\mathbb{Z}_\ell^\times$, and conversely for every $t \in T$ there exists a $t_2 \in T_2$ such that $t_2 = \pm t/\sqrt{\det^*(t)}$ (so that again $\Theta_1(t)$ and $\Theta_1(t_2)$ differ by an ℓ-adic unit). Suppose now that $L(T)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$: then the same is true for $L(T_2)$ (as the two Lie algebras coincide), and therefore by property $(**)$ (cf. theorem 8.3.2) we see that $T_2' = T'$ contains $\mathcal{B}_\ell(p, p)$.
Finally, let (t_1, t_2) be in T and suppose that $[t_1], [t_2]$ are multiples of the identity. By construction, there exists a scalar λ and an element $(w_1, w_2) \in T_2$ such that $(t_1, t_2) = \lambda(w_1, w_2)$. As the only multiples of the identity in $\mathrm{SL}_2(\mathbb{F}_\ell)$ are $\pm\,\mathrm{Id}$, changing λ into $-\lambda$ if necessary we can assume that $[w_1] = \mathrm{Id}$. But then the properties of T_2 imply that $[w_2] = \mathrm{Id}$, so $[t_1] = [\lambda \cdot \mathrm{Id}] = [t_2]$. □

Remark 8.3.4. It is clear from this proof that we can assume $\ell > 5$ without loss of generality. Doing so will simplify some of the arguments. Also note that the property $[t_1] = [\mathrm{Id}] \Leftrightarrow [t_2] = [\mathrm{Id}]$ of the group T of theorem 8.3.2 will be clear from its construction, so we will not comment further on it.

Our final objective is to compute, in terms of k, n_1 and n_2, an integer p such that G contains $\mathcal{B}_\ell(p,p)$. This would be immediate if G were a pro-ℓ group, for then we would simply apply theorem 8.3.3 as it is. In general, however, one needs to take into account the structure of $G(\ell)$, and many different possibilities arise, according to the type of $G_1(\ell), G_2(\ell)$ in the Dickson classification.

The problem of studying G is complicated by the many possibilities for the mutual relationship between G, G_1 and G_2. However, in some situations which we now discuss, the two projections G_1 and G_2 behave essentially independently one of the other: in this case the problem is greatly simplified, and it is possible to exhibit an integer p as above without examining too closely the structure of $G(\ell)$. To identify these cases we start with the following easy lemma:

Lemma 8.3.5. *Suppose G contains an element g of the form (x,y), where $[x]$ is trivial and $[y]$ is nontrivial and of order prime to ℓ. The group G contains an element of the form $(1,z)$, where the order of $[z]$ is the same as the order of $[y]$.*

Proof. Such an element is given by any limit point of the sequence g^{ℓ^n}. □

The following statement is also easy to check (cf. lemma 8.4.17 below for an analogous, but more complicated case):

Lemma 8.3.6. *Let m be a non-negative integer, u_1, u_2 be ℓ-adic units, g_1 be $\begin{pmatrix} 1 & u_1\ell^m \\ 0 & 1 \end{pmatrix}$, g_2 be $\begin{pmatrix} 1 & 0 \\ u_2\ell^m & 1 \end{pmatrix}$, and G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$. If G contains both g_1 and g_2, then it also contains all of $\mathcal{B}_\ell(2m)$.*

Lemma 8.3.7. *Suppose that G contains an element (a,b) such that a is trivial modulo ℓ and the prime-to-ℓ part of the order of $[b]$ is at least 3. Then G' contains $\mathcal{B}_\ell(4n_1 + 16n_2, 8n_2)$.*

Proof. Note that both the hypothesis and the conclusion are invariant under any change of basis, so we can freely change bases to simplify the calculations.
There exists an integer m such that $[b]^{\ell^m}$ has order prime to ℓ; replacing (a,b) with $(a,b)^{\ell^m}$ allows us to assume that the order of $[b]$ is at least 3. By lemma 8.3.5, G contains an element of the form $(1,b')$, where the order of $[b']$ is the same as the order of $[b]$. We can therefore assume $a = 1$.
By hypothesis, for any g_2 in $\mathcal{B}_\ell(n_2)$ we can find a g_1 such that (g_1, g_2) belongs to G. It follows that for any $g_2 \in \mathcal{B}_\ell(n_2)$ the element

$$(1,b')^{-1}(g_1,g_2)(1,b')(g_1,g_2)^{-1} = (1, (b')^{-1} g_2 b' g_2^{-1})$$

belongs to G. Up to a choice of basis, we can assume that either $b' = \begin{pmatrix} d & 0 \\ 0 & 1/d \end{pmatrix}$ for a certain unit d, or $b' = \begin{pmatrix} c & d\varepsilon \\ d & c \end{pmatrix}$ for certain units c, d and a certain ε such that $[\varepsilon]$ is not a square (for this second case cf. lemma 1.4.7). In the first case, choosing $g_2 = \begin{pmatrix} 1 & \ell^{n_2} \\ 0 & 1 \end{pmatrix}$ shows that

$$\left(1, \begin{pmatrix} 1 & (d^{-2}-1)\ell^{n_2} \\ 0 & 1 \end{pmatrix}\right)$$

belongs to G. Given that d is not congruent to ± 1 modulo ℓ, for otherwise the order of $[b]$ would be 1 or 2, we see that the ℓ-adic valuation of $(d^{-2}-1)\ell^{n_2}$ is exactly n_2. Similarly, G contains $\left(1, \begin{pmatrix} 1 & 0 \\ (d^2-1)\ell^{n_2} & 1 \end{pmatrix}\right)$, and by lemma 8.3.6 this implies that G contains $\{1\} \times \mathcal{B}_\ell(2n_2)$. A similar analysis in the second case (taking g_2 of the form $\begin{pmatrix} 1+e & -\frac{ce(e+2)}{d(e+1)} \\ 0 & \frac{1}{1+e} \end{pmatrix}$, where $v_\ell(e) \geq n_2$) shows that G contains $\{1\} \times \mathcal{B}_\ell(4n_2)$.

Consider now an element $h = (h_1, h_2)$ of G whose first coordinate is $h_1 = \begin{pmatrix} 1 & \frac{1}{\ell^2-1}\ell^{n_1} \\ 0 & 1 \end{pmatrix}$; such an element exists by assumption. Its $\ell(\ell^2-1)$-th power is of the form $h' = \left(\begin{pmatrix} 1 & \ell^{n_1+1} \\ 0 & 1 \end{pmatrix}, h_2'\right)$, where $[h_2'] = [h_2]^{\ell(\ell^2-1)} = [h_2]^{|\mathrm{SL}_2(\mathbb{F}_\ell)|} = [\mathrm{Id}]$. The ℓ^{4n_2-1}-th power of h' (recall that $n_2 > 0$), therefore, is a certain

$$h'' = \left(\begin{pmatrix} 1 & \ell^{n_1+4n_2} \\ 0 & 1 \end{pmatrix}, h_2''\right),$$

where $h_2'' \in \mathcal{B}_\ell(4n_2)$. By what we already saw, G contains $(1, h_2'')$, so G contains

$$h''(1, h_2'')^{-1} = \left(\begin{pmatrix} 1 & \ell^{n_1+4n_2} \\ 0 & 1 \end{pmatrix}, 1\right).$$

The same argument shows that G also contains $\left(\begin{pmatrix} 1 & 0 \\ \ell^{n_1+4n_2} & 1 \end{pmatrix}, 1\right)$, and we finally deduce that G contains $\mathcal{B}_\ell(2n_1+8n_2) \times \{1\}$, hence G' contains $\mathcal{B}_\ell(4n_1+16n_2) \times \mathcal{B}_\ell(8n_2)$. □

Lemma 8.3.8. *Suppose that $L(N(G))$ contains an element of the form $(0,u)$, where u is nonzero modulo ℓ^{s+1} (s a non-negative integer). Then G' contains $\{1\} \times \mathcal{B}_\ell(2s+8n_2)$.*
If $L(N(G))$ contains two elements $(u_1,0)$ and $(0,u_2)$ that are nonzero modulo ℓ^{s+1}, then G' contains $\mathcal{B}_\ell(2s+8n_1) \times \mathcal{B}_\ell(2s+8n_2)$.

Proof. Note first that the Lie algebra $L(N(G))$ is stable under conjugation by G (by the same argument as lemma 1.4.5). The smallest Lie subalgebra of $L(N(G))$ that contains $(0,u)$ and is stable under conjugation by G is $0 \oplus L(u)$, where $L(u)$ is the smallest Lie subalgebra of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ that contains u and is stable under conjugation by G_2. By virtue of lemma 8.2.1, and given that G_2 contains $\mathcal{B}_\ell(n_2)$, the algebra $L(u)$ contains $\ell^{s+4n_2}\mathfrak{sl}_2(\mathbb{Z}_\ell)$. It follows that $L(N(G))$ contains $0 \oplus \ell^{s+4n_2}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, and applying theorem 8.3.3 we deduce that $N(G)'$ (hence also G') contains $\{1\} \times \mathcal{B}_\ell(2s+8n_2)$. The second statement is now immediate. □

We now have three categories of groups for which, given information on $L(G)$, we can deduce information on G:

(A) pro-ℓ groups: by theorem 8.3.3, if $L(G)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$, then G' contains $\mathcal{B}_\ell(2k,2k)$;

(B) groups that contain an element (a,b) such that $[a]$ is trivial and the prime-to-ℓ part of the order of $[b]$ is least 3, because of lemma 8.3.7;

(C) groups satisfying the hypotheses of lemma 8.3.8.

We now start with a general closed subgroup G of $\mathrm{SL}_2(\mathbb{Z}_\ell)^2$ and show that (up to passing to a subgroup of finite, absolutely bounded index) the group G must satisfy one of these three sets of hypotheses.
As already anticipated, we will need a case analysis based on the structure of $G_1(\ell)$ and $G_2(\ell)$. These are subgroup of $\mathrm{SL}_2(\mathbb{F}_\ell)$, and by the Dickson classification we know that any subgroup of $\mathrm{SL}_2(\mathbb{F}_\ell)$ is of one of the following types:

trivial, split Cartan, nonsplit Cartan, normalizer of split Cartan, normalizer of nonsplit Cartan, Borel, exceptional, $\mathrm{SL}_2(\mathbb{F}_\ell)$.

Remark 8.3.9. To be more precise we should rather write 'contained in a split Cartan subgroup', 'contained in a Borel subgroup', etc. The slight abuse of language should not cause any problems.

We call 'type' of G the pair (type of $G_1(\ell)$, type of $G_2(\ell)$). The proof will proceed by analysing all the possibilities for the type of G.
Notice that, if (B) above applies, then (without even using any information on $L(G)$) we know that G contains $\mathcal{B}_\ell(4n_1 + 16n_2, 8n_2)$, and we are done. We can therefore assume that (B) does *not* apply and drastically reduce the number of cases we need to treat, as we now show. Consider the kernel J_2 of the reduction $G(\ell) \to G_1(\ell)$, which we identify to a (normal) subgroup of $G_2(\ell)$. If the prime-to-ℓ part of the order of J_2 is at least 3, then we are in case (B) above: indeed $\mathrm{SL}_2(\mathbb{F}_\ell)$ contains only one element of order 2, namely minus the identity, so if the prime-to-ℓ part of the order of J_2 is at least 3, then J_2 contains an element b whose order has prime-to-ℓ part at least 3, and we are done.
Suppose, on the contrary, that the prime-to-ℓ part of the order of J_2 is at most 2. Taking into account that J_2 is a subgroup of $\mathrm{SL}_2(\mathbb{F}_\ell)$, we see that the ℓ-part of its order can either be 1 or ℓ. Thus the order of J_2 can only be $1, 2, \ell, 2\ell$; furthermore, the last two cases can only happen if $G_2(\ell)$ is of Borel type, since these are the only subgroups of $\mathrm{SL}_2(\mathbb{F}_\ell)$ admitting a normal ℓ-Sylow. The same argument also applies to $J_1 = \ker(G(\ell) \to G_2(\ell))$.
Replacing G with its subgroup $H = \ker(G \to G(\ell) \to G_2(\ell) \to G_2(\ell)/N(G_2(\ell)))$ we can assume that the order of J_2 is either 1 or ℓ. Similarly, up to passing to a second subgroup of index 2, we can assume that the order of J_1 is 1 or ℓ.
Goursat's lemma implies that $G_1(\ell)/J_1$ and $G_2(\ell)/J_2$ are isomorphic, and since J_i is either trivial or agrees with $N(G_i(\ell))$ we see that in particular we can assume

$$G(\ell)/N(G(\ell)) \text{ is the graph of an isomorphism between the finite groups } G_1(\ell)/N(G_1(\ell)) \text{ and } G_2(\ell)/N(G_2(\ell)). \qquad (*)$$

In order to minimize the number of cases we need to treat, let us also get rid of the 'exceptional' case and simplify the Cartan ones.
In case $G_1(\ell)$ is exceptional, there exists a subgroup $H < G$ of index dividing 48 with the property that $\pi_1(H)(\ell)$ is a cyclic subgroup of order 3 or 5 (according to whether the projective image of $G_1(\ell)$ is isomorphic to A_4, or respectively to A_5 or S_5). Assumption $(*)$ implies that $\pi_2(H)$ is also cyclic of order either 3 or 5.

Likewise, if $G_1(\ell)$ is contained in the normalizer of a Cartan subgroup, passing to a subgroup H of index at most 2 allows us to assume that $\pi_1(H)(\ell)$ is in fact contained in the Cartan subgroup itself, and the same then holds for $\pi_2(H)(\ell)$.
As a final simplifying assumption, note that if $G_1(\ell) \cong G_2(\ell)$ is of order dividing 8, then passing to the subgroup defined by $\ker(G \to G(\ell))$ ensures that G is in fact pro-ℓ (so (A) applies, and we are done). Notice that this step is not necessary if we performed the reduction from the exceptional to the cyclic case.
Putting it all together, we see that at the cost of passing to a subgroup of index dividing $2 \cdot 2 \cdot 48$ we have reduced the list of our cases (for a single factor) to

$$\text{trivial, split Cartan, nonsplit Cartan, Borel, } \mathrm{SL}_2(\mathbb{F}_\ell),$$

and we can furthermore assume that property $(*)$ holds and that the orders of $G_1(\ell)$ and $G_2(\ell)$ do not divide 8. We shall now list all the remaining cases for the type of G, and then show how to deal with each of them in turn.

Before getting started with our case analysis let us record a couple of simple results on Teichmüller lifts:

Definition 8.3.10. Let F be a finite unramified extension of $\mathbb{Q}_\ell$ of degree k, with residue field $\mathbb{F} = \mathbb{F}_{\ell^k}$. For an element $[f] \in \mathbb{F}$ we denote $\omega([f])$ the Teichmüller lift of $[f]$, that is to say the only element $g \in \mathcal{O}_F$ that reduces to $[f]$ in $\mathbb{F}$ and satisfies $g^{\ell^k} = g$.

Lemma 8.3.11. *With the notation of the previous definition, the sequence $f^{\ell^{kn}}$ converges to $\omega([f])$ when n tends to infinity.*

Proof. Immediate. □

Lemma 8.3.12. *Let g be an element of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ such that $[g]$ has order prime to ℓ and strictly greater than 2. Then the sequence $g^{\ell^{2n}}$ for $n \in \mathbb{N}$ converges to a certain g_∞ that satisfies $g_\infty^{\ell^2} = g_\infty$. Moreover, if $[g]$ is diagonalizable over $\mathbb{F}_\ell$, the limit g_∞ even satisfies $g_\infty^{\ell} = g_\infty$.*

Proof. The assumption implies that $[g]$ has distinct eigenvalues, hence there exists a quadratic unramified extension F of $\mathbb{Q}_\ell$ over which g can be written as $g = P^{-1}DP$, where $D = \mathrm{diag}(\lambda, \lambda^{-1})$ is diagonal and P is a base-change matrix. By the previous lemma, the sequence $D^{\ell^{2n}}$ converges to $\mathrm{diag}(\omega([\lambda]), \omega([\lambda^{-1}]))$, so the sequence $g^{\ell^{2n}} = PD^{\ell^{2n}}P^{-1}$ converges to $P^{-1}\,\mathrm{diag}(\omega([\lambda]), \omega([\lambda^{-1}]))P$, which satisfies the conclusion since $\omega([\lambda])^{\ell^2} = \omega([\lambda])$. For the second statement we can take $F = \mathbb{Q}_\ell$ and use the fact that the Teichmüller lifts satisfy $\omega([\lambda])^\ell = \omega([\lambda])$. □

We are now ready to deal with the remaining cases for the type of G. Every case is dealt with in a separate section, whose title is of the form (type of G_1, type of G_2). Note that since $N(\mathrm{SL}_2(\mathbb{F}_\ell))$ is trivial we have $G_1(\ell) = \mathrm{SL}_2(\mathbb{F}_\ell)$ if and only if $G_2(\ell) = \mathrm{SL}_2(\mathbb{F}_\ell)$; this helps exclude a few more cases. Also notice that thus far we have replaced G by a subgroup of index dividing $4 \cdot 48$, and therefore we are still free, if necessary, to further replace it with subgroups of index dividing 12 (in fact, we shall only need one more replacement, by a subgroup of index 2, in sections 8.3.3 and 8.3.4).

8.3.1 Case (Trivial, anything)

Since we assume $(*)$ and $G_1(\ell)$ is trivial, $G_2(\ell)/N(G_2(\ell))$ is again trivial, so $G_2(\ell)$ is pro-ℓ. The same then holds for G, and therefore G falls into category (A). Theorem 8.3.3 gives directly that G' contains $\mathcal{B}_\ell(2k, 2k)$.

8.3.2 Cases (Nonsplit Cartan, Split Cartan) and (Nonsplit Cartan, Borel)

The same idea works in both cases: consider an element $(a, b) \in G$ with $[a]$ of maximal order in $G_1(\ell)$; notice in particular that $[a]$ has order dividing $\ell + 1$. The finite group $G(\ell)$ contains the element

$$([a], [b])^{\ell(\ell-1)} = ([a]^{\ell(\ell-1)}, [\mathrm{Id}]),$$

and the order of $[a]^{\ell(\ell-1)}$ is given by $\dfrac{\mathrm{ord}([a])}{(\ell - 1, \mathrm{ord}([a]))} =: m$. If m is at least 3 this falls into category (B), contradiction. On the other hand, notice that $(\ell - 1, \mathrm{ord}([a])) \leq 2$, so $m \leq 2$ implies $\mathrm{ord}[a] \mid 4$, contradicting our assumption that the order of $G_1(\ell)$ does not divide 8.

8.3.3 Cases (Split Cartan, Split Cartan), (Borel, Borel) and (Split Cartan, Borel)

We start by considering the type (Split Cartan, Split Cartan), the other two cases being essentially identical. Note that the groups $N(G(\ell)), N(G_1(\ell))$ and $N(G_2(\ell))$ are all trivial, so $G_1(\ell)$ and $G_2(\ell)$ are isomorphic by $(*)$, and we can find an element $(h_1, h_2) \in G$ such that $[h_1], [h_2]$ generate $G_1(\ell), G_2(\ell)$ respectively. By lemma 8.3.12, the limit (g_1, g_2) of the sequence $(h_1, h_2)^{\ell^{2n}}$ satisfies $g_1^\ell = g_1, g_2^\ell = g_2$, and furthermore $[g_1]$, $[g_2]$ generate $G_1(\ell)$, $G_2(\ell)$ respectively. We choose bases in such a way that both g_1 and g_2 are diagonal. Write $g_i = \begin{pmatrix} d_i & 0 \\ 0 & d_i^{-1} \end{pmatrix}$, where d_i satisfies $d_i^\ell = d_i$. Since we are assuming that $G_1(\ell) \cong G_2(\ell)$, we know that the orders of $[d_1]$ and $[d_2]$ agree. In particular, we can write $[d_2] = [d_1]^q$ for a certain integer q, $1 \leq q \leq \mathrm{ord}[d_1]$, that is prime to the order of $[d_1]$. Replacing G with a subgroup of index 2 if necessary we can assume $q \neq \pm 2$.
Given that the Teichmüller lift is a homomorphism we deduce $d_2 = \omega([d_2]) = \omega([d_1])^q = d_1^q$. The cases $q = \pm 1$ and $q \neq \pm 1$ will turn out to be somewhat different, as we will see shortly. Note that the cases $q = 1$ and $q = -1$ are the same up to a change of basis (the one exchanging the two coordinates on the second factor $\mathrm{SL}_2(\mathbb{Z}_\ell)$), so in the scenario $q = \pm 1$ we can in fact assume without loss of generality that $q = 1$. Consider now the three matrices

$$M_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and let π_1 (resp. π_2, π_3) be the linear maps $\mathfrak{sl}_2(\mathbb{Z}_\ell) \to \mathbb{Z}_\ell \cdot M_i$ giving the projection of an element on its M_1 (resp. M_2, M_3) component. A $\mathbb{Z}_\ell$-basis of the Lie algebra $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$ is given by $(M_i \oplus 0), (0 \oplus M_j)$ for $i = 1, 2, 3$ and $j = 1, 2, 3$. Note that both $L(G)$ and $L(N(G))$ are stable under conjugation by (g_1, g_2) (cf. lemma 1.4.5).

Writing elements of $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$ as the 6-dimensional vectors of their coordinates in the basis just described, the action of conjugating by (g_1, g_2) is given by

$$(x_1, x_2, x_3, x_4, x_5, x_6) \mapsto (d_1^2 x_1, x_2, d_1^{-2} x_3, d_2^2 x_4, x_5, d_2^{-2} x_6).$$

In particular, if we denote C the linear operator (acting on $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$)

$$(x, y) \mapsto (g_1 x g_1^{-1}, g_2 y g_2^{-1}),$$

we have

$$\frac{1}{\ell - 1} \sum_{i=0}^{\ell-2} C^i(x_1, x_2, x_3, x_4, x_5, x_6) = (0, x_2, 0, 0, x_4, 0)\,,$$

since $\dfrac{1}{\ell-1} \displaystyle\sum_{i=0}^{\ell-2} d_1^{2i} = \frac{1}{\ell-1} \frac{d_1^{2(\ell-1)} - 1}{d_1^2 - 1} = 0$ (recall that $d_1^\ell = d_1$ and $d_1^2 \neq 1$), and similarly for d_1^{-2} and $d_2^{\pm 2}$. It follows that if $L(G)$ or $L(N(G))$ contains the vector $(x_1, \ldots, x_6)$, then it also contains the vectors $(x_1, 0, x_3, x_4, 0, x_6)$ and $(d_1^{2j} x_1, 0, d_1^{-2j} x_3, d_2^{2j} x_4, 0, d_2^{-2j} x_6)$ (for every integer j). Consider the matrix

$$V = \begin{pmatrix} 1 & 1 & 1 & 1 \\ d_1^2 & d_1^{-2} & d_2^2 & d_2^{-2} \\ d_1^4 & d_1^{-4} & d_2^4 & d_2^{-4} \\ d_1^6 & d_1^{-6} & d_2^6 & d_2^{-6} \end{pmatrix}.$$

This is a Vandermonde matrix, so its determinant does not vanish modulo ℓ as long as $d_1^2 \not\equiv d_1^{-2}$ and $d_1^2 \not\equiv d_2^{\pm 2} \pmod{\ell}$. Recall that we have already assumed that the order of d_1 does not divide 4, so the first condition is automatically satisfied. If $d_1^2 \not\equiv d_2^{\pm 2} \pmod{\ell}$, then, this matrix is invertible in $\mathbb{Z}_\ell$, that is to say the standard basis vectors $(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)$ can be written as linear combinations of the rows of V. In turn, this implies that the vectors

$$(x_1, 0, 0, 0, 0, 0), (0, 0, x_3, 0, 0, 0), (0, 0, 0, x_4, 0, 0), (0, 0, 0, 0, 0, x_6)$$

can be written as $\mathbb{Z}_\ell$-combinations of the four vectors $C^j(x_1, 0, x_3, x_4, 0, x_6)$ for $j = 0, 1, 2, 3$. Equivalently, we have shown that if $q \neq \pm 1$ the Lie algebras $L(G), L(N(G))$ are stable under the projection operators $\pi_1 \oplus 0, \pi_3 \oplus 0, 0 \oplus \pi_1, 0 \oplus \pi_3$.
On the other hand, under our assumptions if $d_1^2 \equiv d_2^{\pm 2} \pmod{\ell}$ then we have $q = 1$ (so $g_1 = g_2$), and an even easier computation shows that $L(G), L(N(G))$ are stable under the projection operators $\pi_1 \oplus \pi_1$ and $\pi_3 \oplus \pi_3$.

Regarding (with a little abuse of notation) the π_i's as maps from $\mathfrak{gl}_2(\mathbb{Z}_\ell)$ to itself, we can write $\Theta_1 = \pi_1 + \pi_2 + \pi_3$, so that $\pi_i \Theta_1 = \pi_i$ for $i = 1, 2, 3$. Further we have the immediate identity

$$M_1 g_i = \begin{pmatrix} 0 & d_i^{-1} \\ 0 & 0 \end{pmatrix} = d_i^{-1} M_1 \quad \text{for } i = 1, 2,$$

whence for any $A \in \mathrm{GL}_2(\mathbb{Z}_\ell)$ we have

$$\pi_1 \Theta_1(A g_i) = \pi_1(A g_i) = d_i^{-1} \pi_1(A) \quad \text{for } i = 1, 2.$$

Again we need to distinguish between the case when d_1^2 and $d_2^{\pm 2}$ are not congruent modulo ℓ and the case when they are. Consider the former. We know that $(0 \oplus \pi_1)(L(G))$ is contained in $L(G)$, and it is generated by an element of the form $(0 \oplus \pi_1)(\Theta_2(h))$, for a certain $h \in G$. We can certainly

choose an integer m in such a way that $h \cdot (g_1, g_2)^m$ belongs to $N(G)$. For such an m, the element $\Theta_2(h \cdot (g_1, g_2)^m)$ lies in $L(N(G))$. However, $L(N(G))$ is stable under $0 \oplus \pi_1$, so it also contains

$$\begin{aligned} d_2^m \cdot (0 \oplus \pi_1)\left(\Theta_2(h \cdot (g_1, g_2)^m)\right) &= d_2^m (0, \pi_1 \Theta_1(h g_2^m)) \\ &= (0, \pi_1(\Theta_1(h))). \end{aligned}$$

Now $L(G)$ contains $(0 \oplus \pi_1)(\ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell)) = 0 \oplus \ell^k M_1$, so the previous formula shows that the same holds for $L(N(G))$. Repeating the same argument swapping the roles of two factors then shows that $L(N(G))$ contains $\ell^k M_1 \oplus 0$. We are in situation (C), and lemma 8.3.8 implies that G' contains

$$\mathcal{B}_\ell\left(2k + 8\max\{n_1, n_2\}, 2k + 8\max\{n_1, n_2\}\right).$$

Suppose on the other hand that $d_1^2 \equiv d_2^{\pm 2} \pmod{\ell}$, so that under our assumptions we have $d_1 \equiv d_2 \pmod{\ell}$. Let us write, for the sake of simplicity, g for $g_1 = g_2$ and d for $d_1 = d_2$. As we have seen, both $L(G)$ and $L(N(G))$, thought of as subsets of $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$, are stable under the maps $\pi_i \oplus \pi_i$ for $i = 1, 2, 3$. Hence $L(G)$ is the direct sum of three rank-2 subalgebras $R_i = (\pi_i \oplus \pi_i)(L(G))$, $i = 1, 2, 3$, with R_i open in $\mathbb{Z}_\ell M_i \oplus \mathbb{Z}_\ell M_i$; similarly, $L(N(G))$ is the direct sum of three algebras $S_i = (\pi_i \oplus \pi_i)(L(N(G)))$, with S_i open in $\mathbb{Z}_\ell M_i \oplus \mathbb{Z}_\ell M_i$. We claim that $S_1 = R_1$. If R_1 is generated by the two elements $(\pi_1 \oplus \pi_1)(\Theta_2(h_1)), (\pi_1 \oplus \pi_1)(\Theta_2(h_2))$, then we can find integers m_1, m_2 such that $h_1(g, g)^{m_1}$ and $h_2(g, g)^{m_2}$ belong to $N(G)$. It follows that for $i = 1, 2$ the algebra S_1 contains

$$d^{m_i}(\pi_1 \oplus \pi_1)(\Theta_2(h_i(g, g)^{m_i})) = d^{m_i} d^{-m_i}(\pi_1 \oplus \pi_1)(\Theta_2(h_i)),$$

i.e. $S_1 = R_1$ as claimed.
Now note that $L(G)$ contains $\ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell)$, so $R_1 = (\pi_1 \oplus \pi_1)(L(G))$ contains $\ell^k M_1 \oplus \ell^k M_1$ and the same is true for $S_1 = (\pi_1 \oplus \pi_1)(L(N(G)))$. As above, we conclude that G' contains

$$\mathcal{B}_\ell\left(2k + 8\max\{n_1, n_2\}, 2k + 8\max\{n_1, n_2\}\right).$$

Finally, note that the (Borel, Borel) and (Split Cartan, Borel) cases are completely analogous: we simply need to choose for (g_1, g_2) a generator of $G/N(G)$, which is cyclic (cf. lemma 1.3.18).

8.3.4 Case (Nonsplit Cartan, Nonsplit Cartan)

We follow an approach very close to that of the previous section. Using lemma 8.3.12 and the fact that $G(\ell)$ is the graph of an isomorphism $G_1(\ell) \to G_2(\ell)$, we can find an element (g_1, g_2) of G such that $g_i^{\ell^2} = g_i$ and $[g_i]$ generates $G_i(\ell)$; in a suitable basis we can write $g_i = \begin{pmatrix} a_i & b_i\varepsilon_i \\ b_i & a_i \end{pmatrix}$, where ε_i is an element of $\mathbb{Z}_\ell^\times \setminus \mathbb{Z}_\ell^{\times 2}$ (that is to say, $[\varepsilon_i]$ is not a square in $\mathbb{F}_\ell^\times$). The condition that the order of $G_i(\ell)$ does not divide 8 implies $a_i b_i \not\equiv 0 \pmod{\ell}$. For any ℓ-adic unit ε consider now the three matrices

$$M_1(\varepsilon) = \begin{pmatrix} 0 & \varepsilon \\ 1 & 0 \end{pmatrix}, \; M_2(\varepsilon) = M_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \; M_3(\varepsilon) = \begin{pmatrix} 0 & -\varepsilon \\ 1 & 0 \end{pmatrix}.$$

A basis of $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$ is given by

$$(M_1(\varepsilon_1), 0), \; (M_2(\varepsilon_1), 0), \; (M_3(\varepsilon_1), 0), \; (0, M_1(\varepsilon_2)), \; (0, M_2(\varepsilon_2)), \; (0, M_3(\varepsilon_2)),$$

and again we write elements of $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$ as six-dimensional vectors in this basis. Let C be the linear operator (from $\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \mathfrak{sl}_2(\mathbb{Z}_\ell)$ to itself) given by $(x, y) \mapsto (g_1, g_2)(x, y)(g_1, g_2)^{-1}$; once again, $L(G)$ and $L(N(G))$ are stable under C. The matrix of C in this basis is block-diagonal, the blocks being given by $B_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 + 2\varepsilon_i b_i^2 & 2a_i b_i \varepsilon_i \\ 0 & 2a_i b_i & 1 + 2\varepsilon_i b_i^2 \end{pmatrix}$. Since the bottom-right 2 by 2 block of B_i is simply g_i^2, the eigenvalues of B_i are 1 and the squares of the eigenvalues of g_i. The analogue of the condition $[d_1]^2 \neq [d_2]^{\pm 2}$ of the previous paragraph is 'the only eigenvalue shared by $[B_1] \in \mathrm{GL}_3(\mathbb{F}_\ell)$ and $[B_2] \in \mathrm{GL}_3(\mathbb{F}_\ell)$ is 1'. We now replace G by a subgroup of index at most 2 by the following prescription (notice that 1 is not an eigenvalue of $[g_i]^2$):

1. for an element $z \in \mathbb{Z}_\ell[\sqrt{\varepsilon_1}, \sqrt{\varepsilon_2}]$ denote by $[z]$ its image in $\overline{\mathbb{F}_\ell}$. By construction we have $a_i \pm b_i\sqrt{\varepsilon_i} = \omega\left([a_i \pm b_i\sqrt{\varepsilon_i}]\right)$, where ω is the Teichmüller lift. If $[a_1 \pm \sqrt{\varepsilon_1}b_1]^2 = [a_2 \pm \sqrt{\varepsilon_2}b_2]^2$, then (if necessary) we apply on the second factor $\mathrm{SL}_2(\mathbb{Z}_\ell)$ the change of basis induced by the matrix $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ to assume $[a_1 + \sqrt{\varepsilon_1}b_1]^2 = [a_2 + \sqrt{\varepsilon_2}b_2]^2$. Notice that the matrix S does not belong to $\mathrm{SL}_2(\mathbb{Z}_\ell)$, but nonetheless both the hypothesis and the conclusion of theorem 8.3.2 are left unchanged by this change of basis. We then set $T := \ker\left(G \to G(\ell) \to G(\ell)/2G(\ell)\right)$, which (since $G(\ell)$ is cyclic) has index 2 in G. The element $(g_1, g_2)^2 \in T$ projects to a generator of $T(\ell)$, and we have
$$(a_1 + b_1\sqrt{\varepsilon_1})^2 = \omega\left([a_1 + b_1\sqrt{\varepsilon_1}]^2\right) = \omega\left([a_2 + b_2\sqrt{\varepsilon_2}]^2\right) = (a_2 + b_2\sqrt{\varepsilon_2})^2,$$
which – using the fact that $a_i^2 - \varepsilon_i b_i^2 = \det g_i = 1$ – implies $\varepsilon_1 b_1^2 = \varepsilon_2 b_2^2$. Notice that T, being of index 2 in G, is normal, hence its Lie algebra $L(T)$ is stable not just under conjugation by elements of T, but also under conjugation by elements of G; the same is true for $L(N(T)) = L(N(G))$. In particular, both these algebra are stable under conjugation by (g_1, g_2), that is, they are stable under C.

2. if $(a_1 \pm \sqrt{\varepsilon_1}b_1)^2$, $(a_2 \pm \sqrt{\varepsilon_2}b_2)^2$ are all distinct in $\overline{\mathbb{F}_\ell}$ we simply set $T = G$. In this case the squares of the eigenvalues of g_1 and of g_2 are distinct.

We now assume that $L(T)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$, and we shall show that T' contains $\mathcal{B}_\ell(2k + 8\max\{n_1, n_2\}, 2k + 8\max\{n_1, n_2\})$.

Suppose first that we are in subcase (2). Let $p_1(x)$ be the characteristic polynomial of B_1, and consider $p_1(C)$. This will be a block-diagonal operator whose first block is the null matrix and whose second block is of the form $\left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{A}} \\ 0 & & \end{array}\right)$, with A invertible modulo ℓ (this follows at once from the fact that the reduction modulo ℓ of this block can be computed as $p_1([B_2])$, and the only eigenvalue that is common to $[B_1]$ and $[B_2]$ is 1). Note furthermore that by the Hamilton-Cayley theorem the 2×2 identity can be expressed as a polynomial in A, so that ultimately the diagonal matrix with diagonal entries $(0, 0, 0, 0, 1, 1)$ can be expressed a polynomial in C. Concretely, this is

the operator

$$\Pi : \left(\begin{pmatrix} h_1 & x_1 \\ y_1 & -h_1 \end{pmatrix}, \begin{pmatrix} h_2 & x_2 \\ y_2 & -h_2 \end{pmatrix}\right) \mapsto \left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} h_2 & \frac{x_2 - \varepsilon_2 y_2}{2} \\ \frac{\varepsilon_2 y_2 - x_2}{2\varepsilon_2} & -h_2 \end{pmatrix}\right),$$

and we have just shown that $L(T)$ and $L(N(T))$ (being stable under C) are in particular also stable under Π. As T_2 contains $\mathcal{B}_\ell(n_2)$, there exists an element f_1 of T_1 such that $\left(f_1, \begin{pmatrix} 1 & \ell^{n_2} \\ 0 & 1 \end{pmatrix}\right)$ belongs to T. Taking the $\ell(\ell^2-1)$-th power of this element shows that $N(T)$ contains the element

$$\left(f_1^{\ell(\ell^2-1)}, \begin{pmatrix} 1 & (\ell^2-1)\ell^{n_2+1} \\ 0 & 1 \end{pmatrix}\right),$$

and therefore $L(N(T))$ contains

$$\frac{1}{\ell^2-1}\Theta_2\left(f_1^{\ell(\ell^2-1)}, \begin{pmatrix} 1 & (\ell^2-1)\ell^{n_2+1} \\ 0 & 1 \end{pmatrix}\right) = \left(\frac{1}{\ell^2-1}\Theta_1\left(f_1^{\ell(\ell^2-1)}\right), \begin{pmatrix} 0 & \ell^{n_2+1} \\ 0 & 0 \end{pmatrix}\right).$$

Applying Π and multiplying by $2\varepsilon_2$ we see that $L(N(T))$ contains $\left(0, \begin{pmatrix} 0 & \varepsilon_2\ell^{n_2+1} \\ -\ell^{n_2+1} & 0 \end{pmatrix}\right)$; by lemma 1.4.8 we see that $L(N(T))$ also contains $\left(0, \begin{pmatrix} \ell^{n_2+1} & 0 \\ 0 & -\ell^{n_2+1} \end{pmatrix}\right)$, and since $L(N(T))$ is stable under conjugation by all of T it is easy to see that $L(N(T))$ contains $0 \oplus \ell^{2n_2+1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Swapping the roles of T_1, T_2 and repeating the same argument we find that $L(N(T))$ contains all of $\ell^{2n_1+1}\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^{2n_2+1}\mathfrak{sl}_2(\mathbb{Z}_\ell)$, and applying theorem 8.3.3 we deduce that $N(T)'$ (hence T') contains $\mathcal{B}_\ell(4n_1+2, 4n_2+2)$.

Next consider subcase (1). Recall that the algebras $L(T)$ and $L(N(T))$ are stable under C. We keep the notation $M_i(\varepsilon)$ from subcase (2), and we let $\pi_1(\varepsilon)$ (resp. $\pi_2(\varepsilon), \pi_3(\varepsilon)$) be the linear maps $\mathfrak{sl}_2(\mathbb{Z}_\ell) \to \mathbb{Z}_\ell \cdot M_i(\varepsilon)$ giving the projection of an element on its $M_1(\varepsilon)$ (resp. $M_2(\varepsilon), M_3(\varepsilon)$) component. Using the fact that $\varepsilon_1 b_1^2 = \varepsilon_2 b_2^2$, one easily checks that

$$\pi_1(\varepsilon_1) \oplus \pi_1(\varepsilon_2) = -\frac{1}{4\varepsilon_1 b_1^2}\left(\mathrm{Id} - 2(1 + 2\varepsilon_1 b_1^2)C + C^2\right),$$

from which we see that $L(T), L(N(T))$ are stable under $\pi_1(\varepsilon_1) \oplus \pi_1(\varepsilon_2)$ and therefore, by difference, also under $\tilde{\pi} : (x_1, x_2, x_3, x_4, x_5, x_6) \mapsto (0, x_2, x_3, 0, x_5, x_6)$. We now set Λ to be the 6×6 matrix of C in the basis $(M_1(\varepsilon_1), 0), (M_2(\varepsilon_1), 0), (M_3(\varepsilon_1), 0), (0, M_1(\varepsilon_2)), (0, M_2(\varepsilon_2)), (0, M_3(\varepsilon_2))$; as we have already seen in subcase (2), this is the block-diagonal operator with blocks given by $1, g_1^2, 1, g_2^2$. We claim that $\tilde{\pi}(L(T))$ and $\tilde{\pi}(L(N(T)))$, seen as submodules of $\mathbb{Z}_\ell^6$, are stable under left multiplication by Λ^{-1}. Indeed, since the Lie algebras of T and of $N(T)$ are stable under conjugation by (g_1, g_2) the claim follows from the identity

$$\tilde{\pi}\left((g_1, g_2)^{-1}(t_1, t_2)(g_1, g_2)\right) = \Lambda^{-1} \cdot \tilde{\pi}\left((t_1, t_2)\right) \quad \forall (t_1, t_2) \in \mathfrak{sl}_2(\mathbb{Z}_\ell)^2. \tag{8.1}$$

Furthermore, one easily checks that, for all $t \in T$, we have $\tilde{\pi}\left(\Theta_2\left((g_1, g_2)^2 \cdot t\right)\right) = \Lambda \cdot \tilde{\pi}\left(\Theta_2(t)\right)$. Let now $w_1, \ldots, w_4 \in T$ be such that $\tilde{\pi}(L(T))$ is generated by $\tilde{\pi}(\Theta_2(w_1)), \ldots, \tilde{\pi}(\Theta_2(w_4))$. Since $[(g_1^2, g_2^2)]$ generates $T(\ell)$, for $i = 1, \ldots, 4$ there is an integer m_i such that $(g_1, g_2)^{m_i} w_i$ belongs to $N(T)$ (that is, it is trivial modulo ℓ): it follows that $\Theta_2\left((g_1^2, g_2^2)^{m_i} w_i\right)$ is in $L(N(T))$, and since $L(N(T))$ is

stable under both $\tilde{\pi}$ and Λ^{-1} we find

$$\Lambda^{-m_i} \cdot \tilde{\pi}\left(\Theta_2\left((g_1^2, g_2^2)^{m_i} w_i\right)\right) = \Lambda^{-m_i} \cdot \Lambda^{m_i} \cdot \tilde{\pi}(\Theta_2(w_i)) = \tilde{\pi}\left(\Theta_2(w_i)\right) \in L(N(T)).$$

This easily implies $L(N(T)) \supseteq \tilde{\pi}\left(L(N(T))\right) = \tilde{\pi}(L(T)) \supseteq \tilde{\pi}(\ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k \mathfrak{sl}_2(\mathbb{Z}_\ell))$. In particular, $L(N(T))$ contains two elements $(u_1, 0)$ and $(0, u_2)$ with $u_1, u_2 \not\equiv 0 \pmod{\ell^{k+1}}$: by lemma 8.3.8 we conclude that $N(T)'$ contains $\mathcal{B}_\ell\left(2k + 8\max\{n_1, n_2\}, 2k + 8\max\{n_1, n_2\}\right)$.

8.3.5 Case $(\mathrm{SL}_2(\mathbb{F}_\ell), \mathrm{SL}_2(\mathbb{F}_\ell))$

We reduce this case to the question of whether or not, for any given t, the group $G(\ell^t)$ is the graph of an isomorphism $G_1(\ell^t) \to G_2(\ell^t)$. The following lemma covers the case when this does *not* happen:

Lemma 8.3.13. *Let m be a positive integer. Suppose G contains an element of the form (g_1, g_2), where g_1 is trivial modulo ℓ^m but g_2 is not. Then G' contains $\mathcal{B}_\ell(4m, 4m)$.*

Proof. Let

$$x_1 = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}, y_1 = \begin{pmatrix} 1 & 0 \\ \ell & 1 \end{pmatrix}, h_1 = \begin{pmatrix} 1+\ell & 0 \\ 0 & \frac{1}{1+\ell} \end{pmatrix}.$$

As G_1 is all of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ (cf. lemma 1.3.15) we can certainly find $x_2, y_2, h_2 \in G_2$ such that $x = (x_1, y_1)$, $y = (y_2, y_2)$ and $h = (h_1, h_2)$ belong to G. Recall that $G(\ell)$ is the graph of an isomorphism $G_1(\ell) \to G_2(\ell)$: as x_1, y_1 and h_1 are all trivial modulo ℓ, the same must be true of x_2, y_2, h_2. Consider then the elements $x^{\ell^{m-1}}, y^{\ell^{m-1}}$ and $h^{\ell^{m-1}}$. They satisfy:

- their first coordinates generate $\mathcal{B}_\ell(m)$
- their second coordinates are trivial modulo ℓ^m,

so that the group they generate contains an element of the form (g_1^{-1}, g_2'), where g_2' is necessarily trivial modulo ℓ^m. The group G, therefore, contains the product $g = (g_1^{-1}, g_2')(g_1, g_2) = (1, g_2' g_2)$, whose second coordinate is congruent to g_2 (and therefore nontrivial) modulo ℓ^m. Notice that by assumption $G(\ell)$ is the graph of an isomorphism $G_1(\ell) \to G_2(\ell)$, so since g_1 is trivial modulo ℓ the same is true for g_2; it follows in particular that $g \in N(G) = \ker(G \to G(\ell))$. Thus $L(N(G))$ contains $\Theta_2(g)$, which is of the form $(0, u)$ with u nontrivial modulo ℓ^m. Applying lemma 8.3.8 we deduce that $N(G)$ contains $\{1\} \times \mathcal{B}_\ell(2m)$ (notice that in the notation of lemma 8.3.8 we can take $s = 0$). To finish the proof, consider the group H topologically generated by $x' = x^{\ell^{2m-1}}, y' = y^{\ell^{2m-1}}, h' = h^{\ell^{2m-1}}$. It is clear that $\pi_1(H) \supseteq \mathcal{B}_\ell(2m)$ and $\pi_2(H) \subseteq \mathcal{B}_\ell(2m)$, so the group generated by H and $\{1\} \times \mathcal{B}_\ell(2m)$ (which is still a subgroup of G) contains $\mathcal{B}_\ell(2m) \times \mathcal{B}_\ell(2m)$, and we are done. □

We now show that for $t = k + 1$ the hypothesis of the previous lemma is satisfied. Indeed suppose by contradiction that the projections

$$G(\ell^{k+1}) \to G_1(\ell^{k+1}), \quad G(\ell^{k+1}) \to G_2(\ell^{k+1})$$

have trivial kernel. Then Goursat's lemma implies that $G(\ell^{k+1})$ is the graph of an isomorphism $G_1(\ell^{k+1}) \to G_2(\ell^{k+1})$, i.e. an automorphism of $\mathrm{SL}_2(\mathbb{Z}/\ell^{k+1}\mathbb{Z})$. By [146, Theorem 2], and since $\ell > 5$, all such automorphisms are inner, so we can find a matrix $M \in \mathrm{SL}_2(\mathbb{Z}/\ell^{k+1}\mathbb{Z})$ such that

$$G(\ell^k) = \left\{(x, y) \in \mathrm{SL}_2(\mathbb{Z}/\ell^{k+1}\mathbb{Z})^2 \;\middle|\; y = MxM^{-1}\right\}.$$

Consequently, if we still denote by the same letter any lift of M to $\mathrm{SL}_2(\mathbb{Z}_\ell)$, we have

$$G \subseteq \left\{(x,y) \in \mathrm{SL}_2(\mathbb{Z}_\ell)^2 \;\middle|\; y \equiv MxM^{-1} \pmod{\ell^{k+1}}\right\}.$$

Applying Θ_2 and noticing that $\mathrm{tr}(MxM^{-1}) = \mathrm{tr}(x)$ we deduce

$$L(G) \subseteq \left\{(x,y) \in \mathfrak{sl}_2(\mathbb{Z}_\ell)^2 \;\middle|\; y \equiv MxM^{-1} \pmod{\ell^{k+1}}\right\},$$

but this contradicts the hypothesis that $L(G)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Thus at least one of the two projections $G(\ell^{k+1}) \to G_i(\ell^{k+1})$ has nontrivial kernel, and the previous lemma shows that G' contains $\mathcal{B}_\ell(4(k+1), 4(k+1))$.

8.4 $\ell = 2$, $n = 2$

In this section we prove:

Theorem 8.4.1. *Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_2) \times \mathrm{GL}_2(\mathbb{Z}_2)$ whose projection modulo 4 is trivial. Denote by G_1, G_2 the two projections of G on the factors $\mathrm{GL}_2(\mathbb{Z}_2)$, and let $n_1 \geq 4, n_2 \geq 4$ be integers such that G_i contains $\mathcal{B}_2(n_i)$. Suppose furthermore that for every $(g_1, g_2) \in G$ we have $\det(g_1) = \det(g_2) \in 1 + 8\mathbb{Z}_2$. If $L(G)$ contains $2^k\mathfrak{sl}_2(\mathbb{Z}_2) \oplus 2^k\mathfrak{sl}_2(\mathbb{Z}_2)$ for a certain $k \geq 2$, then G contains*

$$\mathcal{B}_2(12(k + 12n_2 + 5n_1 + 13) + 1, 12(k + 12n_1 + 5n_2 + 13) + 1).$$

By an argument similar to that used for the case of odd ℓ (and that will be carried out at the end of this section) we can easily reduce the problem to one concerning subgroups of $\mathrm{SL}_2(\mathbb{Z}_2)$:

Theorem 8.4.2. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)^2$ whose reduction modulo 4 is trivial. Denote by G_1, G_2 the two projections of G on the factors $\mathrm{SL}_2(\mathbb{Z}_2)$, and choose integers $n_i \geq 4$ so that G_i contains $\mathcal{B}_2(n_i)$. If $L(G)$ contains $\ell^k\mathfrak{sl}(\mathbb{Z}_2) \oplus \ell^k\mathfrak{sl}(\mathbb{Z}_2)$ for a certain integer $k \geq 2$, then G contains all of*

$$\mathcal{B}_2(6(k + 12n_2 + 5n_1 + 13), 6(k + 12n_1 + 5n_2 + 13)).$$

The proof of this theorem, although technically involved, relies on a very simple idea: we can find an element of G of the form (Id, a), where a is not too close to the identity 2-adically, and this easily implies the conclusion by theorem 1.5.2. In order to find a we proceed by contradiction: if there is no such a, then G looks very much like the graph of a map $G_1 \to G_2$, and this imposes severe restrictions on its Lie algebra. Quantifying this idea of 'being 2-adically very close to a graph' gives a contradiction with the fact that $L(G)$ contains $\ell^k\mathfrak{sl}(\mathbb{Z}_2) \oplus \ell^k\mathfrak{sl}(\mathbb{Z}_2)$.

We start to deploy the strategy just described by showing that it is in fact enough to find an element (Id, a) as above:

Lemma 8.4.3. *Suppose that G contains an element of the form (Id, a), where a is nontrivial modulo 2^n for a certain integer $n \geq 3$. Then G contains all of*

$$\mathcal{B}_2(6n + 24n_2 + n_1 + 24, 6n + 24n_2 + 24).$$

Proof. Consider the smallest normal subgroup H of G that contains (Id, a). This is clearly of the form $\{\mathrm{Id}\} \times H_2$, where H_2 is the smallest normal subgroup of G_2 containing a. The Lie algebra L of H_2 contains $\Theta_1(a)$, so it is nontrivial modulo 2^n. By normality of H_2 in G_2, L is stable under conjugation by $\mathcal{B}_2(n_2)$, so lemma 8.2.1 says that L contains $2^{n+4n_2+4}\mathfrak{sl}_2(\mathbb{Z}_2)$. Applying theorem 1.5.2 we deduce that H_2 contains $\mathcal{B}_2(6n + 24n_2 + 24)$. We finish the proof as we did for lemma 8.3.13. □

We now come to the hard part of the proof, namely showing that the non-existence of such an a implies that G is very close to being a graph. For any fixed integer $t > 2$, we distinguish two possibilities:

1. There exist two elements (a, b) and (a, b') of G with $b \not\equiv b' \pmod{2^t}$, or equivalently, there exists an element of G of the form (Id, b'') with $b'' \not\equiv \mathrm{Id} \pmod{2^t}$. In this case we simply apply lemma 8.4.3.

2. For every $a \in G_1$ there exists a (necessarily unique) $b \in G_2(2^t)$ such that, for every element of G of the form (a, c), we have $c \equiv b \pmod{2^t}$. In this case we write $b = \varphi(a)$, so that φ is a well-defined function $G_1 \to G_2(2^t)$.

As it is clear, the key step in proving theorem 8.4.2 is to bound the values of t for which this second case can arise. Let then $t \geq 3$ be an integer for which we are in case 2. Choose a function $\psi : G_1 \to G_2$ such that

- $\psi(a) \equiv \varphi(a) \pmod{2^t}$ for every $a \in G_1$;
- $(a, \psi(a))$ belongs to G for every $a \in G_1$.

As we shall see shortly, the function φ is actually a continuous group morphism. On the other hand, the function ψ does not necessarily have any nice group-theoretic properties, but allows us to work with well-defined elements of $\mathbb{Z}_2$ instead of congruence classes. We will also see that any such morphism φ is, in a suitable sense, 'inner', a fact that will lead to a contradiction for t large enough. From now on, therefore, we work under the following assumption:

Condition 8.4.4. The integer $t \geq 3$ has the following property: for every $a \in G_1$ there exists a (necessarily unique) $b \in G_2(2^t)$ such that, for every element of G of the form (a, c), we have

$$c \equiv b \pmod{2^t}.$$

Lemma 8.4.5. *φ defines a group morphism $G_1 \to G_2(2^t)$.*

Proof. Let a_1, a_2 be any two elements of G_1. Then $(a_1, \psi(a_1))(a_2, \psi(a_2)) = (a_1a_2, \psi(a_1)\psi(a_2))$ belongs to G, so our assumption implies that $\psi(a_1)\psi(a_2) \equiv \varphi(a_1a_2) \pmod{2^t}$. As $\psi(a_1)$ (resp. $\psi(a_2)$) is congruent to $\varphi(a_1)$ (resp. $\varphi(a_2)$) modulo 2^t the claim follows. □

Definition 8.4.6. For any integer $n \geq 2$ we let

$$x(n) = \begin{pmatrix} 1 & 2^n \\ 0 & 1 \end{pmatrix}, \quad y(n) = \begin{pmatrix} 1 & 0 \\ 2^n & 1 \end{pmatrix}, \quad h(n) = \begin{pmatrix} 1+2^n & 0 \\ 0 & \frac{1}{1+2^n} \end{pmatrix}.$$

To ease the notation, for $i = 1, 2$ we also let x_i (resp. y_i, h_i) denote $x(n_i)$ (resp. $y(n_i), h(n_i)$).

Recall that, by assumption, G_1 contains $\mathcal{B}_2(n_1)$, hence it contains x_1, y_1, h_1.

Lemma 8.4.7. *φ is continuous.*

Proof. Denote by $\pi_1, \pi_2 : \mathrm{SL}_2(\mathbb{Z}_2)^2 \to \mathrm{SL}_2(\mathbb{Z}_2)$ the projections on the two factors. As x_1, y_1, h_1 belong to G_1 we can find a, b, c so that $x' = (x_1, a), y' = (y_1, b), h' = (h_1, c)$ all belong to G. Consider then $(x')^{2^t}, (y')^{2^t}, (h')^{2^t}$ and the group H they generate (topologically). The projection $\pi_1(H)$ contains $x_1^{2^t} = x(n_1 + t)$ and $y_1^{2^t} = y(n_1 + t)$, hence it contains $\mathcal{B}_2(2(n_1 + t))$ and is therefore open in $\mathrm{SL}_2(\mathbb{Z}_2)$. On the other hand, $\pi_2(H)$ is generated by $a^{2^t}, b^{2^t}, c^{2^t}$, so it is trivial modulo 2^t. It follows that for any $g_1 \in H$ we have $\varphi(g_1) = 1$, so $\ker \varphi$ is open and φ is continuous. □

Definition 8.4.8. Let g be an element of $\mathrm{SL}_2(\mathbb{Z}_2)$ (resp. of a finite quotient $\mathrm{SL}_2(\mathbb{Z}/2^m\mathbb{Z})$) that is trivial modulo 2, and let β be any 2-adic integer. Write $\beta = \sum_{n\geq 0} a_n 2^n$, where each a_n is either 0 or 1. We set $g^\beta = \prod_{n\geq 0} g^{a_n 2^n}$, which is well-defined since for every finite p only a finite number of terms appearing in the product are nontrivial modulo 2^p.

Lemma 8.4.9. *Let β be any 2-adic integer and g be an element of G_1. We have $\varphi(g^\beta) = \varphi(g)^\beta$.*

Proof. Write $\beta = \sum_{n\geq 0} a_n 2^n$ (with $a_n \in \{0, 1\}$), and set $\beta(s) = \sum_{n\leq s} a_n 2^n$, so that $\beta(s)$ is a rational integer for all s. In view of the continuity of φ, the lemma follows from

$$\varphi\left(g^\beta\right) = \varphi\left(\lim_{s\to\infty} g^{\beta(s)}\right) = \lim_{s\to\infty} \varphi\left(g^{\beta(s)}\right) = \lim_{s\to\infty} \varphi\left(g\right)^{\beta(s)} = \varphi(g)^\beta.$$

□

Let, for the sake of simplicity, $\alpha = (1 + 2^{n_1})^2$. Note that $h_1 x_1 h_1^{-1} = x_1^\alpha$, so – by the previous lemma – we have $\varphi(h_1)\varphi(x_1)\varphi(h_1)^{-1} = \varphi(x_1)^\alpha$, or equivalently

$$\psi(h_1)\psi(x_1)\psi(h_1)^{-1} \equiv \psi(x_1)^\alpha \pmod{2^t}.$$

Taking the logarithm of both sides we deduce

$$\psi(h_1)\log(\psi(x_1))\psi(h_1)^{-1} \equiv \alpha \log \psi(x_1) \pmod{2^t}. \tag{8.2}$$

Lemma 8.4.10. *Suppose that $\log \psi(x_1)$ vanishes modulo $2^{n_1+n_2}$. Then G contains*

$$\mathcal{B}_2(30n_1 + 30, 30n_1 + n_2 + 30).$$

Proof. Exponentiating the hypothesis gives $\psi(x_1) \equiv \mathrm{Id} \pmod{2^{n_1+n_2}}$. There exist $a, b, c \in G_1$ such that $x' = (a, x_2), y' = (b, y_2), h' = (c, h_2)$ belong to G. Consider $(x')^{2^{n_1}}, (y')^{2^{n_1}}, (h')^{2^{n_1}}$: these three elements generate a group H such that $\pi_1(H)$ is trivial modulo 2^{n_1+1} (recall that a, b, c are already trivial modulo 4) and $\pi_2(H)$ contains $\mathcal{B}_2(n_1 + n_2)$. It follows that H (hence G) contains an element of the form $(w, \psi(x_1)^{-1})$, where w is trivial modulo 2^{n_1+1}. Therefore G contains the element $(x_1, \psi(x_1))(w, \psi(x_1)^{-1}) = (x_1 w, 1)$, where $x_1 w \equiv x_1 \pmod{2^{n_1+1}}$ is nontrivial modulo 2^{n_1+1}. The claim follows from lemma 8.4.3. □

Lemma 8.4.11. *With the notation of theorem 8.4.2, condition 8.4.4 and definition 8.4.8, let*

$$U = t - 3n_1 - n_2 - 4$$

and suppose that $\log \psi(x_1)$ does not vanish modulo $2^{n_1+n_2}$. Suppose furthermore that $U > 3n_1$. Then $\psi(h_1)$ is diagonalizable (over $\mathbb{Q}_2$), with eigenvalues λ_1, λ_2 that satisfy

$$\lambda_1 \equiv 1 + 2^{n_1} \pmod{2^U}, \quad \lambda_2 \equiv (1 + 2^{n_1})^{-1} \pmod{2^U}.$$

Proof. Denote $\mathcal{C}_{\psi(h_1)}$ the linear endomorphism of $\mathfrak{sl}_2(\mathbb{Z}_2)$ given by conjugation by $\psi(h_1)$ and $p(x)$ its characteristic polynomial. Note that $\mathrm{tr}(\log\psi(x_1)) = \log\det\psi(x_1) = 0$, so $\log\psi(x_1)$ is in $\mathfrak{sl}_2(\mathbb{Z}_2)$. Finally, let λ_1, λ_2 be the eigenvalues of $\psi(h_1)$.

An easy computation shows that $p(x) = (x-1)\left(x-\lambda_1^2\right)\left(x-\lambda_2^2\right)$. With a little abuse of notation, in the course of the proof we shall use congruences (modulo powers of 2) that involve λ_1, λ_2: a priori, these might not be elements of $\mathbb{Z}_2$, so the precise meaning of these congruences is that we work with the ideals generated by the relevant powers of 2 in the ring of integers of F, where F is a suitable quadratic extension of $\mathbb{Q}_2$ that contains λ_1, λ_2.

Since the logarithm map commutes with conjugation by $\psi(h_1)$, from equation (8.2) we get

$$\begin{aligned}\psi(h_1)\left(\log\psi(x_1)\right)\psi(h_1)^{-1} &= \log\left(\psi(h_1)\psi(x_1)\psi(h_1)^{-1}\right)\\ &= \log\left(\psi(x_1)^\alpha + O(2^t)\right)\\ &= \alpha\log\psi(x_1) + O(2^t),\end{aligned}$$

and therefore $\log\psi(x_1)$ is an approximate eigenvector for $\mathcal{C}_{\psi(h_1)}$. We deduce from lemma 8.2.2 and the assumption $\log\psi(h_1) \not\equiv 0 \pmod{2^{n_1+n_2}}$ that $p(\alpha) \equiv 0 \pmod{2^{t-n_1-n_2}}$. We have $\lambda_1 \equiv \lambda_2 \equiv 1 \pmod 4$ by construction, so $v_2(1+2^{n_1}+\lambda_i) = 1$ for $i = 1, 2$. Hence $v_2(p(\alpha))$, which is given by

$$v_2(\alpha-1) + v_2(1+2^{n_1}+\lambda_1) + v_2(1+2^{n_1}+\lambda_2) + v_2(1+2^{n_1}-\lambda_1) + v_2(1+2^{n_1}-\lambda_2),$$

does not exceed

$$(n_1+1) + 1 + 1 + 2\max_i v_2(1+2^{n_1}-\lambda_i),$$

so that

$$\max_i v_2(1+2^{n_1}-\lambda_i) \geq \frac{v_2(p(\alpha)) - n_1 - 3}{2} \geq \frac{t-2n_1-n_2-3}{2}.$$

Let $U' = \left\lfloor \dfrac{t-2n_1-n_2-3}{2} \right\rfloor$. For $U' > 2n_1$ (a condition that is implied by the hypothesis $U > 3n_1$) we have $\lambda_1 \equiv 1+2^{n_1} \pmod{2^{U'}}$ and $\lambda_2 \equiv \lambda_1^{-1} \equiv 1-2^{n_1}+2^{2n_1} \pmod{2^{2n_1+1}}$. It follows in particular that $v_2(1+2^{n_1}-\lambda_2) = n_1+1$, so that we can improve our previous estimate to

$$v_2(1+2^{n_1}-\lambda_1) \geq v_2(p(\alpha)) - (n_1+1) - 1 - 1 - (n_1+1) \geq t - 3n_1 - n_2 - 4.$$

If we let $U = t-3n_1-n_2-4$, this amounts to saying that $\lambda_1 \equiv \lambda_2^{-1} \equiv 1+2^{n_1} \pmod{2^U}$. Note that the trace of $\psi(h_1)$ is given by

$$\lambda_1 + \lambda_2 = 1 + 2^{n_1} + 1 - 2^{n_1} + 2^{2n_1} + O(2^{3n_1}),$$

at least for $U > 3n_1$, so

$$\mathrm{tr}(\psi(h_1))^2 - 4\det(\psi(h_1)) = \left(2 + 2^{2n_1} + O(2^{3n_1})\right)^2 - 4 = 2^{2n_1+2} + O(2^{3n_1+1})$$

is a square in $\mathbb{Z}_2$ (since $n_1 \geq 4$). It follows that the eigenvalues of $\psi(h_1)$ lie in $\mathbb{Z}_2$, because

$$\lambda_{1,2} = \frac{\mathrm{tr}(\psi(h_1)) \pm \sqrt{(\mathrm{tr}\,\psi(h_1))^2 - 4}}{2}$$

is in $\mathbb{Q}_2$ (as the expression under square root is a square) and is 2-integral (as $p(x)$ is monic with 2-integral coefficients). It follows that $\psi(h_1)$ is diagonalizable, and that its eigenvalues satisfy the given congruences. □

Corollary 8.4.12. *Under the hypotheses of the previous lemma, there exists a* 2*-integral matrix* $N \in \mathrm{GL}_2(\mathbb{Q}_2)$ *that satisfies:*

1. $N^{-1}\psi(h_1)N$ *is diagonal (with diagonal entries* λ_1, λ_2 *as above);*
2. $v_2 \det(N) \leq n_1 + 1$.

Proof. Let w_1, w_2 be two eigenvectors for $\psi(h_1)$, associated resp. with λ_1, λ_2, and chosen so as to be 2-integral and to have at least one coordinate that is a 2-adic unit. Let N be the matrix having w_1, w_2 as columns: it is clear that N satisfies (1). Now if w_1, w_2 are linearly independent over $\mathbb{F}_2$ we are done, for then $v_2(\det N) = 0$. Otherwise, up to rescaling w_1, w_2 and swapping the two coordinates, we can assume they are of the form $w_1 = (1, w_1')^T, w_2 = (1, w_2')^T$. The determinant of N is simply $w_2' - w_1'$, so that we have

$$\begin{pmatrix} 0 \\ w_2' - w_1' \end{pmatrix} \equiv 0 \pmod{\det(N)} \quad \Rightarrow \quad w_2 \equiv w_1 \pmod{\det(N)}.$$

Applying $\psi(h_1)$ to both sides of this last congruence we find

$$\lambda_2 w_2 = \lambda_1 w_1 \pmod{\det(N)},$$

and comparing the first coordinates of these vectors we deduce $\lambda_1 \equiv \lambda_2 \pmod{\det(N)}$. Since $\lambda_1 \equiv 1 + 2^{n_1} \pmod{2^{2n_1}}, \lambda_2 \equiv 1 - 2^{n_1} \pmod{2^{2n_1}}$, we have in particular $2^{n_1+1} \equiv 0 \pmod{\det(N)}$, whence the corollary. □

Assuming the hypotheses of lemma 8.4.11, fix a matrix N as in the previous corollary, and change basis on the second factor $\mathrm{SL}_2(\mathbb{Z}_2) \subseteq \mathrm{SL}_2(\mathbb{Q}_2)$ using N. As it is clear, in this basis there is no guarantee that the elements of G_2 are 2-integral. We restrict our attention to those that are:

Lemma 8.4.13. *Asume that* $\log \psi(x_1)$ *does not vanish modulo* $2^{n_1+n_2}$ *and that* $U > 3n_1$*, so that we can find an* N *as above. Let* g_1 *be an element of* $\mathcal{B}_2(2n_1 + 1) \subseteq G_1$*. Then* $N^{-1}\psi(g_1)N$ *is* 2*-integral and trivial modulo 4.*

Proof. As $\mathcal{B}_2(2n_1+1)$ is generated by $x(2n_1+1), y(2n_1+1), h(2n_1+1)$ it is enough to show the lemma for these three elements. Let us only do the first, the proof being virtually identical for the other two. We have $x(2n_1 + 1) = x(n_1)^{2^{n_1+1}}$, so $\psi(x(2n_1 + 1)) \equiv \psi(x(n_1))^{2^{n_1+1}} \pmod{2^t}$. As $\psi(x(n_1))$ is trivial modulo 4, the matrix $\psi(x(n_1))^{2^{n_1+1}}$ is trivial modulo 2^{n_1+3}. Writing $\psi(x(n_1))^{2^{n_1+1}}$ as $\mathrm{Id} + 2^{n_1+3}B$ for a certain 2-integral matrix B we have

$$N^{-1}\psi(x(2n_1 - 1))N = N^{-1}\left(\mathrm{Id} + 2^{n_1+3}B\right)N = \mathrm{Id} + N^*\left(\frac{2^{n_1+3}}{\det(N)}B\right)N,$$

where $N^* = \det(N)N^{-1}$ is the adjugate matrix of N. This last expression is manifestly 2-integral and congruent modulo 4 to the identity. □

Let N^* be the adjugate matrix of N and $D = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = N^{-1}\psi(h_1)N$. By the previous lemma, the following identity only involves 2-integral matrices:

$$\begin{aligned}\left(N^{-1}\psi(h_1)N\right)\left(N^{-1}\psi(x(2n_1+1))N\right)\left(N^*\psi(h_1)^{-1}N\right) &= N^*\psi(h_1)\psi(x(2n_1+1))\psi(h_1)^{-1}N \\ &\equiv N^*\varphi(h_1)\varphi(x(2n_1+1))\varphi(h_1)^{-1}N \pmod{2^t} \\ &\equiv N^*\varphi\left(x(2n_1+1)^\alpha\right)N \pmod{2^t} \\ &\equiv N^*\varphi\left(x(2n_1+1)\right)^\alpha N \pmod{2^t}.\end{aligned}$$

Dividing through by $\det(N)$ we deduce

$$D\left(N^{-1}\psi(x(2n_1+1))N\right)D^{-1} \equiv \left(N^{-1}\psi(x(2n_1+1))N\right)^\alpha \pmod{2^{t-n_1-1}}. \tag{8.3}$$

Note that since $t-n_1-1 \geq U$ we can in particular rewrite this last equation modulo 2^U instead, in which case we also know that $D \equiv \begin{pmatrix} 1+2^{n_1} & \\ & (1+2^{n_1})^{-1} \end{pmatrix} \pmod{2^U}$. Furthermore, we see that

$$\begin{aligned}v_2(\alpha-\lambda_2^2) &= v_2\left((1+2^{n_1})^2 - \left(1-2^{n_1}+O\left(2^{2n_1}\right)\right)^2\right) \\ &= v_2\left(1+2^{n_1+1}+2^{2n_1} - \left(1-2^{n_1+1}+O\left(2^{2n_1}\right)\right)\right) \\ &= v_2\left(2^{n_1+2}+O(2^{2n_1})\right) \\ &= n_1+2,\end{aligned}$$

and similarly $v_2(\alpha-1) = n_1+1$.

Lemma 8.4.14. *Let* $A_x = \log\left(N^{-1}\psi(x(2n_1+1))N\right)$. *Write*

$$A_x = \mu_x\begin{pmatrix}0&1\\0&0\end{pmatrix} + \mu_y\begin{pmatrix}0&0\\1&0\end{pmatrix} + \mu_h\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

for certain scalars μ_x, μ_y, μ_h. *We have*

$$\mu_h \equiv 0 \pmod{2^{U-n_1-1}}, \quad \mu_y \equiv 0 \pmod{2^{U-n_1-2}}.$$

Proof. Reducing equation (8.3) modulo 2^U and taking logarithms we get

$$\alpha(A_x) \equiv DA_xD^{-1} \equiv \begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix} A_x \begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix}^{-1} \pmod{2^U},$$

and the right hand side can be computed explicitly in terms of μ_x, μ_h, μ_y. We arrive at

$$\alpha A_x \equiv \lambda_1^2\mu_x\begin{pmatrix}0&1\\0&0\end{pmatrix} + \lambda_2^2\mu_y\begin{pmatrix}0&0\\1&0\end{pmatrix} + \mu_h\begin{pmatrix}1&0\\0&-1\end{pmatrix} \pmod{2^U},$$

i.e.

$$\begin{cases}\alpha\mu_x \equiv \lambda_1^2\mu_x \pmod{2^U} \\ \alpha\mu_y \equiv \lambda_2^2\mu_y \pmod{2^U} \\ \alpha\mu_h \equiv \mu_h \pmod{2^U}.\end{cases}$$

Rewriting the last formula as $(\alpha-1)\mu_h \equiv 0 \pmod{2^U}$ shows that $\mu_h \equiv 0 \pmod{2^{U-n_1-1}}$, while the second congruence guarantees $v_2(\mu_y) + v_2(\alpha-\lambda_2^2) \geq U$, whence $v_2(\mu_y) \geq U-n_1-2$. □

A completely analogous argument yields similar congruences for $\log\left(N^{-1}\psi(y(2n_1+1))N\right)$. Now a simple series computation, using the fact that $\begin{pmatrix}0&1\\0&0\end{pmatrix}$ squares to zero, shows that

$$N^{-1}\psi(x(2n+1))N = \exp\log\left(N^{-1}\psi(x(2n+1))N\right) = \exp(A_x) = \mathrm{Id} + A_x + O(2^{U-2n_1}),$$

and since we have similar expressions for $\psi(y(2n+1))$ we arrive at:

Proposition 8.4.15. *Assume that*

- $T := U - 2n_1 = t - 5n_1 - n_2 - 4$ *is larger than* $3n_1$*;*
- G *contains no element of the form* (Id, b)*, where* $b \not\equiv \mathrm{Id} \pmod{2^t}$*;*
- $\log\psi(x_1)$ *does not vanish modulo* $2^{n_1+n_2}$*.*

There exists a 2-integral matrix $N \in \mathrm{GL}_2(\mathbb{Q}_2)$*, whose determinant satisfies* $v_2\det(N) \leq n_1+1$*, and scalars* $c, d \in 4\mathbb{Z}_2$*, such that*

$$N^{-1}\psi(h_1)N \equiv \begin{pmatrix}\lambda_1 & \\ & \lambda_2\end{pmatrix} \equiv h_1 \pmod{2^T},$$

$$N^{-1}\psi(x(2n_1+1))N \equiv \begin{pmatrix}1&c\\0&1\end{pmatrix} \pmod{2^T}, \quad N^{-1}\psi(y(2n_1+1))N \equiv \begin{pmatrix}1&0\\d&1\end{pmatrix} \pmod{2^T}.$$

Remark 8.4.16. As we shall see shortly, the product cd is 2-adically very close to 2^{4n_1+2}, as one would expect. However, it is not true in general that c, d, taken separately, are 2-adically very close to 2^{2n_1+1}.

The parameters c, d are up to now completely free, and they can't be controlled in any way by simply using the relations $hxh^{-1} = x^\alpha$, $hyh^{-1} = y^{-\alpha}$ (which are just the integrated forms of the usual $\mathfrak{sl}_2$-Lie algebra relations $[h,x] = 2x, [h,y] = -2y$). In order to say something meaningful about them, we shall need to use an integrated form of the Lie algebra relation $[x,y] = h$, that is to say we want to have some degree of control on the commutator $xyx^{-1}y^{-1}$. This is made possible by the following simple lemma, whose proof is immediate by induction:

Lemma 8.4.17. *For every* $a \in \mathbb{Z}_2$ *of valuation at least 1 set*

$$x_a = \begin{pmatrix}1&a\\0&1\end{pmatrix}, y_a = \begin{pmatrix}1&0\\a&1\end{pmatrix}.$$

Then

- *For any pair* (a,b) *of elements of* $\mathbb{Z}_2$ *of valuation at least 1, the finite products*

$$\Pi_n = \prod_{i=-n}^{-1} x_a^{(ab)^{-i}} \cdot \left(x_a y_b x_a^{-1} y_b^{-1}\right) \cdot \prod_{i=1}^{n} y_b^{-(ab)^i}$$

converge, as $n \to \infty$*, to* $\begin{pmatrix}\frac{1}{1-ab}&0\\0&1-ab\end{pmatrix}$.

- *Let (a,b) be as above and (c,d) be any other pair of elements of 2-adic valuation at least 1. The finite products*

$$\Pi'_n = \prod_{i=-n}^{-1} x_c^{(ab)^{-i}} \cdot \left(x_c y_d x_c^{-1} y_d^{-1}\right) \cdot \prod_{i=1}^{n} y_d^{-(ab)^i}$$

converge to a limit for $n \to \infty$, and this limit is of the form $\begin{pmatrix} \star & \star \\ \star & 1-cd \end{pmatrix}$.

Apply this lemma to $a = 2^{2n_1+1}, b = 2^{2n_1+1}$: the infinite product

$$\prod_{i=-\infty}^{-1} x_a^{(ab)^{-i}} \cdot \left(x_a y_b x_a^{-1} y_b^{-1}\right) \cdot \prod_{i=1}^{\infty} y_b^{-(ab)^i}$$

converges to $\begin{pmatrix} \frac{1}{1-2^{4n_1+2}} & 0 \\ 0 & 1-2^{4n_1+2} \end{pmatrix} = h_1^\beta$, where β is defined by $(1+2^{n_1})^{-\beta} = 1 - 2^{4n_1+2}$. Applying φ (which, being continuous, commutes with infinite products) we deduce that

$$\varphi(h_1)^\beta = \prod_{i=-\infty}^{-1} \varphi(x_a)^{(ab)^{-i}} \cdot \left(\varphi(x_a)\varphi(y_b)\varphi(x_a)^{-1}\varphi(y_b)^{-1}\right) \cdot \prod_{i=1}^{\infty} \varphi(y_b)^{-(ab)^i}.$$

Set, for the sake of notational simplicity, $B_x = N^{-1}\psi(x(2n_1+1))N, B_y = N^{-1}\psi(y(2n_1+1))N$. Multiplying by N^*, N and dividing by $\det(N)$ (as we did for example in deriving equation (8.3)) we get

$$\left(N^{-1}\varphi(h_1)N\right)^\beta \equiv \prod_{i=-\infty}^{-1} B_x^{(ab)^{-i}} \cdot \left(B_x B_y B_x^{-1} B_y^{-1}\right) \cdot \prod_{i=1}^{\infty} B_y^{-(ab)^i} \pmod{2^{T-n_1-1}}.$$

Letting c, d be elements of $\mathbb{Z}_2$ that satisfy

$$B_x = N^{-1}\psi(x(2n_1+1))N \equiv \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \pmod{2^T}$$

and

$$B_y = N^{-1}\psi(y(2n_1+1))N \equiv \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix} \pmod{2^T}$$

and applying the second part of the previous lemma to x_c, y_d we obtain

$$(N^{-1}\varphi(h_1)N)^\beta \equiv \prod_{i=-\infty}^{-1} x_c^{(ab)^{-i}} \cdot \left(x_c y_d x_c^{-1} y_d^{-1}\right) \cdot \prod_{i=1}^{\infty} y_d^{-(ab)^i} = \begin{pmatrix} \star & \star \\ \star & 1-cd \end{pmatrix} \pmod{2^{T-n_1-1}},$$

so that, comparing the bottom-right coefficients, we deduce $1 - 2^{4n_1+2} \equiv 1 - cd \pmod{2^{T-n_1-1}}$. In particular, if $T \geq 5n_1 + 4$, we must have $v_2(c) + v_2(d) = 4n_1 + 2$, and by symmetry we can assume that $v_2(c) \leq 2n_1 + 1$. We deduce that $d \equiv \frac{2^{4n_1+2}}{c} \pmod{2^{T-n_1-1-v_2(c)}}$, and therefore

$$d \equiv \frac{2^{4n_1+2}}{c} \pmod{2^{T-3n_1-2}}.$$

Consider $M = \begin{pmatrix} 1 & 0 \\ 0 & 2^{2n_1+1}/c \end{pmatrix}$ (which, by our assumption on c, is 2-integral). By construction it satisfies $Mx(2n_1+1) = x_c M$, so that

$$Mx(2n_1+1) \equiv x_c M \equiv N^{-1}\psi(x(2n_1+1))NM \pmod{2^T},$$

and furthermore (since $N^{-1}\psi(h_1)N$ is diagonal and congruent to h_1 modulo 2^T) we also have

$$Mh_1 \equiv h_1 M \equiv N^{-1}\psi(h_1)NM \pmod{2^T}.$$

Finally, using what we just proved on d we find (for $T \geq 5n_1 + 4$)

$$\begin{aligned} My(2n_1+1) &= \begin{pmatrix} 1 & 0 \\ 2^{4n_1+2}/c & 2^{2n_1+1}/c \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ d & 2^{2n_1+1}/c \end{pmatrix} = y_d M \\ &\equiv N^{-1}\psi(y(2n+1))NM \pmod{2^{T-3n_1-2}}. \end{aligned}$$

Multiplying any of these equations on the left by $M^* = \begin{pmatrix} 2^{2n_1+1}/c & 0 \\ 0 & 1 \end{pmatrix}$ we get equations of the form $\det M \cdot y(2n_1+1) \equiv M^* N^{-1}\psi(y(2n_1+1))NM \pmod{2^{T-3n_1-2}}$, and similar ones for $x(2n+1), h_1$. Given that $v_2(\det M) \leq 2n_1 + 1$, dividing by $\det M$ we find

$$(NM)^{-1}\psi(y(2n+1))NM \equiv y(2n+1) \pmod{2^{T-5n_1-3}},$$

along with similar relations for $x(2n_1+1), h_1$. As $x(2n_1+1), y(2n_1+1), h_1$ generate $\mathcal{B}_2(2n_1+1)$ we have thus established

Proposition 8.4.18. *For every $g \in \mathcal{B}_2(2n_1+1)$ we have $(NM)^{-1}\psi(g)(NM) \equiv g \pmod{2^{T-5n_1-3}}$.*

We now give a version of the previous proposition that applies to all elements of G_1. Take any element $g \in G_1$. Clearly $g^{2^{2n_1-1}}$ belongs to $\mathcal{B}_2(2n_1+1)$, so

$$(NM)^{-1}\psi\left(g\right)^{2^{2n_1-1}}(NM) \equiv (NM)^{-1}\psi\left(g^{2^{2n_1-1}}\right)(NM) \equiv g^{2^{2n_1-1}} \pmod{2^{T-5n_1-3}}.$$

Notice now that $g, \psi(g)$ are trivial modulo 4 by assumption, so we are allowed to take logarithms, and we obtain

$$2^{2n_1-1}(NM)^{-1}\log\psi(g)(NM) \equiv 2^{2n_1-1}(\log g) \pmod{2^{T-5n_1-3}},$$

whence

$$(NM)^{-1}\log\psi(g)(NM) \equiv \log g \pmod{2^{T-7n_1-2}}.$$

Since $\log g$ is trivial modulo 4, we can exponentiate both sides of the congruence to find

$$(NM)^{-1}\psi(g)(NM) \equiv g \pmod{2^{T-7n_1-2}},$$

a formula which is now valid for every $g \in G_1$. Taking the trace of this last congruence also gives $\mathrm{tr}\,\psi(g) \equiv \mathrm{tr}(g) \pmod{2^{T-7n_1-2}}$. Including again all the assumptions we made along the way, we have thus established:

Proposition 8.4.19. *Assume:*

1. *G contains no element of the form (Id, b), where $b \not\equiv \mathrm{Id} \pmod{2^t}$;*
2. *$\log\psi(x_1)$ does not vanish modulo $2^{n_1+n_2}$;*
3. *$t - 5n_1 - n_2 - 4 \geq 7n_1 + 2$ (so that $T - 7n_1 - 2 \geq 0$).*

For every $g \in G_1$ we have

$$(NM)^{-1}\Theta_1(\psi(g))(NM) \equiv \Theta_1(g) \pmod{2^{T-7n_1-2}}.$$

Notice now that we can replace NM by λNM for any $\lambda \in \mathbb{Z}_2$ and the previous proposition still holds; in particular, we can assume that NM is 2-integral, with at least one coefficient which is a 2-adic unit.

Corollary 8.4.20. *Under the same assumptions, the Lie algebra $L(G)$ is contained in*

$$\left\{((NM)z, z(NM)) \;\middle|\; z \in \mathfrak{sl}_2(\mathbb{Z}_\ell)\right\}$$

when regarded modulo $2^{t-12n_1-n_2-8}$.

The result we were aiming for is now well within reach:

Proof. (of theorem 8.4.2) Suppose that $L(G)$ contains $2^k\mathfrak{sl}_2(\mathbb{Z}_2) \oplus 2^k\mathfrak{sl}_2(\mathbb{Z}_2)$. Since one of the coefficients of NM is a 2-adic unit, this contradicts the conclusion of the previous corollary if $t-12n_1-n_2-8 > k$, so at least one of the assumptions cannot hold if we take $t = k+12n_1+n_2+9$. Now for this choice of t the inequality given in condition 3 is certainly satisfied, so either 1 or 2 must fail. If 2 fails, then lemma 8.4.10 implies that G contains $\mathcal{B}_2(30(n_1+1), 30(n_1+1)+n_2)$. On the other hand, if condition 1 is not met, then lemma 8.4.3 implies that G contains all of

$$\mathcal{B}_2(6(k+12n_1+5n_2+13)+n_1, 6(k+12n_1+5n_2+13)) \subseteq \mathcal{B}_2(30(n_1+1), 30(n_1+1)+n_2).$$

Finally, note that the hypotheses of the theorem are symmetric in n_1, n_2, so we can repeat the whole argument switching the roles of G_1, G_2, which shows that G also contains

$$\mathcal{B}_2(6(k+12n_2+5n_1+13), 6(k+12n_2+5n_1+13)+n_2)$$

and concludes the proof of the theorem. □

As promised, we can finally deduce theorem 8.4.1:

Proof. Let G^{sat} be the group generated by G and by $\mathbb{Z}_2^\times \cdot (\mathrm{Id}, \mathrm{Id})$, denote by U the intersection $G^{\mathrm{sat}} \cap \mathrm{SL}_2(\mathbb{Z}_2)^2$, and let U_1, U_2 be the two projections of U on the factors $\mathrm{SL}_2(\mathbb{Z}_2)$. Note that $U' = G'$: it suffices to show that any element of G, when multiplied by a suitable scalar, lies in U, and this follows from the fact that the determinant of any element of G is a square in $\mathbb{Z}_2^\times$.
Also remark that if G_1 contains $\mathcal{B}_2(n_1)$, then the same is true for U_1: indeed for any $g_1 \in \mathcal{B}_2(n_1)$ we know that there exists a certain $h_2 \in G_2$ such that $(g_1, h_2) \in G$. As $\det(h_2) = \det(g_1) = 1$ by assumption, this shows that (g_1, h_2) belongs to U as well, and therefore g_1 belongs to U_1. The same argument obviously also works for U_2. Applying theorem 8.4.2 to U we deduce that U contains $\mathcal{B}_2(6(k+12n_2+5n_1+13), 6(k+12n_1+5n_2+13))$, and therefore $G' = U'$ contains $\mathcal{B}_2(12(k+12n_2+5n_1+13)+1, 12(k+12n_1+5n_2+13)+1)$ as claimed. □

8.5 Conclusion of the proof

We are now in a position to show that cases $n = 1, 2$ of theorem 8.1.3 (in the form given in sections 8.3 and 8.4) imply the general one. Before doing so, let us remark that the condition $[G : H] \leq 120$ appearing in the statement of theorem 1.4.2 can be improved to $[G : H] \mid 24$: this follows immediately from the same proof and the simple remark that if $G(\ell)$ is exceptional, then

G contains a subgroup H, of index dividing 24, such that $H(\ell)$ is cyclic of order dividing 6 or 10. With this small improvement, case $n = 1$ of theorem 8.1.3 is (amply) covered by theorems 1.4.1 and 1.5.1.

Let us start with the case $\ell \neq 2$. If $\ell = 3$ we take $H = \ker(G \to \mathrm{SL}_2(\mathbb{F}_\ell)^n)$, which is a pro-$\ell$ group and for which the claim follows directly from Pink's theorem (theorem 8.3.3).
By our previous remarks we can then assume $\ell \geq 5, n \geq 2$. Denote by $\pi_i : \mathrm{SL}_2(\mathbb{Z}_\ell)^n \to \mathrm{SL}_2(\mathbb{Z}_\ell)$ the n canonical projections and let $\pi_{i_1,i_2} = \pi_{i_1} \times \pi_{i_2}$ be the projection on the two factors numbered i_1 and i_2. For m between 1 and n, construct inductively groups H_m as follows: apply theorem 1.4.1 to $\pi_1(G)$ to find a subgroup K_1 of $\pi_1(G)$ of index dividing 24 and having property $(\star)$, and set $H_1 = \pi_1^{-1}(K_1)$ (this is a subgroup of G of index dividing 24). Assuming we have constructed K_m and H_m, apply theorem 1.4.1 to $\pi_{m+1}(H_m)$ to find a subgroup K_{m+1} of index dividing 24 and having property $(\star)$, and set $H_{m+1} = \pi_{m+1}^{-1}(K_{m+1})$. It is clear by construction that H_n is a closed subgroup of G of index dividing 24^n. For $1 \leq i \leq n$ let n_i be a positive integer such that $\mathcal{B}_\ell(n_i)$ is contained in $\pi_i(H_n)$.
Now let $(i_1, j_1), (i_2, j_2), \ldots, (i_{n(n-1)/2}, j_{n(n-1)/2})$ be the list of all $n(n-1)/2$ pairs $\{i, j\}$ with $i < j$, and construct inductively groups H_{i_k,j_k} (for $k = 1, \ldots, n(n-1)/2$) as follows. Applying theorem 8.3.2 to $\pi_{i_1,j_1}(H_n)$ we see that at least one of the following holds:

1. $\mathcal{B}_\ell(4n_{i_1} + 16n_{j_1}, 8n_{j_1})$ is contained in $\pi_{i_1,j_1}(H_n)$;
2. there exists a closed subgroup K_{i_1,j_1} of $\pi_{i_1,j_1}(H_n)$ satisfying $[\pi_{i_1,j_1}(H_n) : K_{i_1,j_1}] \mid 48^2$ and having property $(**)$.

We set H_{i_1,j_1} to be the inverse image of K_{i_1,j_1} in H_n in case (2), and to be all of H_n in case (1). We now repeat the procedure: if H_{i_k,j_k} has been constructed, we apply theorem 8.3.2 to $\pi_{i_{k+1},j_{k+1}}(H_{i_k,j_k})$ and construct $H_{i_{k+1},j_{k+1}}$ according to the above prescription, that is to say

1. either $\mathcal{B}_\ell(4n_{i_{k+1}} + 16n_{j_{k+1}}, 8n_{j_{k+1}})$ is contained in $\pi_{i_{k+1},j_{k+1}}(H_{i_k,j_k})$, in which case we set $H_{i_{k+1},j_{k+1}} = H_{i_k,j_k}$,
2. or there exists a closed subgroup $K_{i_{k+1},j_{k+1}}$ of $\pi_{i_{k+1},j_{k+1}}(H_{i_k,j_k})$, of index dividing 48^2 and having property $(**)$, in which case we set $H_{i_{k+1},j_{k+1}}$ to be the inverse image in H_{i_k,j_k} of $K_{i_{k+1},j_{k+1}}$.

We finally set $H = H_{i_{n(n-1)/2},j_{n(n-1)/2}}$; by construction, it is a closed subgroup of G of index dividing $24^n 48^{n(n-1)}$.
Denote by $\tau_i : \mathfrak{sl}_2(\mathbb{Z}_\ell)^n \to \mathfrak{sl}_2(\mathbb{Z}_\ell)$ (resp. $\tau_{i,j}$) the projection on the i-th (resp. (i,j)-th) factor. Suppose that $L(H)$ contains $\ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ldots \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$. We have

$$L(K_i) \supseteq L(\pi_i(H)) \supseteq \tau_i(L(H)) \supseteq \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell),$$

so the properties of K_i imply that it contains $\mathcal{B}_\ell(4k)$. Note now that $[H_i : H]$ is only divisible by factors 2 and 3, hence the same is true for $[\pi_i(H_i) : \pi_i(H)]$. As $\mathcal{B}_\ell(4k)$ is a pro-ℓ group, ℓ is neither 2 nor 3, and $\mathcal{B}_\ell(4k) \subseteq \pi_i(H_i)$, it follows that $\mathcal{B}_\ell(4k) \subseteq \pi_i(H)$. In particular, all the integers n_j introduced above can be taken to be $4k$. Consider now a pair of indices (i, j). As before we have $L(K_{i,j}) \supseteq L(\pi_{i,j}(H)) \supseteq \tau_{i,j}(L(H)) \supseteq \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell) \oplus \ell^k\mathfrak{sl}_2(\mathbb{Z}_\ell)$, so two cases arise (depending on whether we were in case (1) or (2) above):

1. either $\mathcal{B}_\ell(4n_i + 16n_j, 8n_j) \supseteq \mathcal{B}_\ell(80k, 32k)$ is contained in $\pi_{i,j}(G)$,

2. or $K_{i,j}$ contains $\mathcal{B}_\ell(p, p)$ with $p = 2k + \max\{2k + 4, 8n_i, 8n_j\} \leq 34k$.

Either way, we see that $\pi_{i,j}(G)$ contains $\mathcal{B}_\ell(80k, 80k)$. Again the only prime factors appearing in $[\pi_{i,j}(G) : \pi_{i,j}(H)]$ are 2 and 3, so the fact that $\pi_{i,j}(G)$ contains $\mathcal{B}_\ell(80k, 80k)$ implies that $\pi_{i,j}(H)$ contains $\mathcal{B}_\ell(80k, 80k)$. Since this holds for every pair of indices $1 \leq i < j \leq n$, lemma 8.2.4 then implies that H contains $\prod_{i=1}^{n} \mathcal{B}_\ell\left(80(n-1)k\right)$, as claimed.

The case $\ell = 2$ is even simpler. Define H to be the kernel of the reduction $G \to \mathrm{SL}_2(\mathbb{Z}/4\mathbb{Z})^n$. Suppose that $L(H)$ contains $2^k\mathfrak{sl}_2(\mathbb{Z}_2) \oplus \ldots \oplus 2^k\mathfrak{sl}_2(\mathbb{Z}_2)$, and let $H_i = \pi_i(H), H_{i,j} = \pi_{i,j}(H)$. Since $L(\pi_i(H)) \supseteq 2^k\mathfrak{sl}_2(\mathbb{Z}_2)$, theorem 1.5.2 implies that H_i contains $\mathcal{B}_2(6k)$, and the integers n_i can all be taken to be $6k > 4$. Similarly, $L(H_{i,j})$ contains $2^k\mathfrak{sl}_2(\mathbb{Z}_2) \oplus 2^k\mathfrak{sl}_2(\mathbb{Z}_2)$, hence by theorem 8.4.2 the group $H_{i,j}$ contains $\mathcal{B}_2(618k + 78, 618k + 78)$: lemma 8.2.4 then implies that H contains $\prod_{i=1}^{n} \mathcal{B}_2\left((n-1)(618k + 79)\right)$. Finally, as H is trivial modulo 4, it is clear that $k \geq 3$, so we have $618k + 79 \leq 645k$ and we are done. □

Bibliographie

[1] A. A. Albert. A solution of the principal problem in the theory of Riemann matrices. *Ann. Math. (2)*, 35(3):500–515, July 1934.

[2] A. A. Albert. On the construction of Riemann matrices. II. *Ann. Math. (2)*, 36(2):376–394, April 1935.

[3] S. Arias-de-Reyna, C. Armana, V. Karemaker, M. Rebolledo, L. Thomas, and N. Vila. Galois representations and Galois groups over Q. *ArXiv e-prints*, July 2014.

[4] E. Artin and J. Tate. *Class field theory*. AMS Chelsea Publishing, Providence, RI, 2009. ISBN 978-0-8218-4426-7. Reprinted with corrections from the 1967 original.

[5] M. Aschbacher. On the maximal subgroups of the finite classical groups. *Invent. Math.*, 76 (3):469–514, 1984. ISSN 0020-9910. doi: 10.1007/BF01388470. URL http://dx.doi.org/10.1007/BF01388470.

[6] G. Banaszak, W. Gajda, and P. Krasoń. On the image of ℓ-adic Galois representations for abelian varieties of type I and II. *Doc. Math., Extra Volume: John H. Coates Sixtieth Birthday*, pages 35–75.

[7] G. Banaszak, W. Gajda, and P. Krasoń. On Galois representations for abelian varieties with complex and real multiplications. *J. Number Theory*, 100(1):117–132, 2003. ISSN 0022-314X. doi: 10.1016/S0022-314X(02)00121-X. URL http://dx.doi.org/10.1016/S0022-314X(02)00121-X.

[8] G. Banaszak, W. Gajda, and P. Krasoń. On the image of Galois ℓ-adic representations for abelian varieties of type III. *Tohoku Mathematical Journal*, 62:163–189, 2010. doi: 10.2748/tmj/1277298644.

[9] D. Bertrand. Galois orbits on abelian varieties and zero estimates. In *Diophantine analysis (Kensington, 1985)*, volume 109 of *London Math. Soc. Lecture Note Ser.*, pages 21–35. Cambridge Univ. Press, Cambridge, 1986.

[10] M. Bhargava and B. H. Gross. The average size of the 2-Selmer group of Jacobians of hyperelliptic curves having a rational Weierstrass point. In *Automorphic representations and L-functions*, volume 22 of *Tata Inst. Fundam. Res. Stud. Math.*, pages 23–91. Tata Inst. Fund. Res., Mumbai, 2013.

[11] C. Birkenhake and H. Lange. *Complex Abelian Varieties*. Springer, 2004. ISBN 3540204881. URL http://www.worldcat.org/isbn/3540204881.

[12] D. M. Bloom. The subgroups of PSL(3, q) for odd q. *Trans. Amer. Math. Soc.*, 127:150–178, 1967. ISSN 0002-9947.

[13] F. A. Bogomolov. Sur l'algébricité des représentations ℓ-adiques. *C.R. Acad. Sci. Paris Sér. A-B*, 290(15):701–703, 1980.

[14] M. V. Borovoĭ. The action of the Galois group on the rational cohomology classes of type (p, p) of abelian varieties. *Mat. Sb. (N.S.)*, 94(136):649–652, 656, 1974.

[15] S. Bosch, W. Lütkebohmert, and M. Raynaud. *Néron models*, volume 21 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1990. ISBN 3-540-50587-3. doi: 10.1007/978-3-642-51438-8. URL http://dx.doi.org/10.1007/978-3-642-51438-8.

[16] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. ISSN 0747-7171. doi: 10.1006/jsco.1996.0125. URL http://dx.doi.org/10.1006/jsco.1996.0125. Computational algebra and number theory (London, 1993).

[17] N. Bourbaki. *Lie groups and Lie algebras. Chapters 4–6*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2002. ISBN 3-540-42650-7. doi: 10.1007/978-3-540-89394-3. URL http://dx.doi.org/10.1007/978-3-540-89394-3. Translated from the 1968 French original by Andrew Pressley.

[18] N. Bourbaki. *Lie Groups and Lie Algebras: Chapters 7-9 (Elements of Mathematics)*. Springer Publishing Company, Incorporated, 2008. ISBN 354068851X, 9783540688518.

[19] J. N. Bray, D. F. Holt, and C. M. Roney-Dougal. *The maximal subgroups of the low-dimensional finite classical groups*, volume 407 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2013. ISBN 978-0-521-13860-4. doi: 10.1017/CBO9781139192576. URL http://dx.doi.org/10.1017/CBO9781139192576. With a foreword by Martin Liebeck.

[20] R. W. Carter. *Finite groups of Lie type*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1985. ISBN 0-471-90554-2. Conjugacy classes and complex characters, A Wiley-Interscience Publication.

[21] W. C. Chi. l-adic and λ-adic representations associated to abelian varieties defined over number fields. *Amer. J. Math.*, 114(2):315–353, 1992. ISSN 0002-9327. doi: 10.2307/2374706. URL http://dx.doi.org/10.2307/2374706.

[22] M. J. Collins. Modular analogues of Jordan's theorem for finite linear groups. *J. Reine Angew. Math.*, 624:143–171, 2008. ISSN 0075-4102. doi: 10.1515/CRELLE.2008.084. URL http://dx.doi.org/10.1515/CRELLE.2008.084.

[23] P. Deligne, J. S. Milne, A. Ogus, and K. Shih. *Hodge cycles, motives, and Shimura varieties*, volume 900 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1982. ISBN 3-540-11174-3.

[24] L. V. Dieulefait. Explicit determination of the images of the Galois representations attached to abelian surfaces with $\mathrm{End}(A) = \mathbb{Z}$. *Experiment. Math.*, 11(4):503–512 (2003), 2002. ISSN 1058-6458. URL http://projecteuclid.org/euclid.em/1057864660.

[25] L. V. Dieulefait and V. Rotger. The arithmetic of QM-abelian surfaces through their Galois representations. *J. Algebra*, 281(1):124–143, 2004. ISSN 0021-8693. doi: 10.1016/j.jalgebra.2004.07.019. URL http://dx.doi.org/10.1016/j.jalgebra.2004.07.019.

[26] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73 (3):349–366, 1983. ISSN 0020-9910. doi: 10.1007/BF01388432. URL http://dx.doi.org/10.1007/BF01388432.

[27] G. Faltings. Complements to Mordell. In *Rational points (Bonn, 1983/1984)*, Aspects Math., E6, pages 203–227. Vieweg, Braunschweig, 1984.

[28] É. Gaudron and G. Rémond. Polarisations et isogénies. *Duke Math. J.*, 163(11):2057–2108, 2014. ISSN 0012-7094. doi: 10.1215/00127094-2782528. URL http://dx.doi.org/10.1215/00127094-2782528.

[29] R. Griess. Schur multipliers of finite simple groups of Lie type. *Trans. Amer. Math. Soc.*, 183:355–421, 1973. ISSN 0002-9947.

[30] A. Grothendieck. Modèles de Néron et monodromie. In *Séminaire de Géométrie Algébrique, Volume 7, Esposé 9*.

[31] A. Grothendieck, M. Raynaud, and D. S. Rim. *Groupes de monodromie en géométrie algébrique. I.* Lecture Notes in Mathematics, Vol. 288. Springer-Verlag, 1972. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 I).

[32] R. Guralnick, M. Larsen, and C. Manack. Low degree representations of simple Lie groups. *Proc. Amer. Math. Soc.*, 140(5):1823–1834, 2012. ISSN 0002-9939. doi: 10.1090/S0002-9939-2011-11007-4. URL http://dx.doi.org/10.1090/S0002-9939-2011-11007-4.

[33] R. M. Guralnick, M. Larsen, and P. H. Tiep. Representation growth in positive characteristic and conjugacy classes of maximal subgroups. *Duke Math. J.*, 161(1):107–137, 2012. ISSN 0012-7094. doi: 10.1215/00127094-1507300. URL http://dx.doi.org/10.1215/00127094-1507300.

[34] C. Hall. An open-image theorem for a general class of abelian varieties. *Bull. Lond. Math. Soc.*, 43(4):703–711, 2011. ISSN 0024-6093. doi: 10.1112/blms/bdr004. URL http://dx.doi.org/10.1112/blms/bdr004. With an appendix by Emmanuel Kowalski.

[35] F. Hazama. Algebraic cycles on certain abelian varieties and powers of special surfaces. *J. Fac. Sci., Univ. Tokyo, Sect. I A*, 31:487–520, 1984.

[36] F. Hazama. Algebraic cycles on nonsimple abelian varieties. *Duke Mathematical Journal*, 58: 31–37, 1989. doi: 10.1215/S0012-7094-89-05803-1.

[37] M. Hindry and A. Pacheco. An analogue of the Brauer-Siegel theorem for abelian varieties in positive characteristic. *Moscow Math. J.*, 2015. To appear. Available at https://sites.google.com/site/amilcarpachecoresearch/publications.

[38] M. Hindry and N. Ratazzi. Torsion dans un produit de courbes elliptiques. *J. Ramanujan Math. Soc.*, 25(1):81–111, 2010. ISSN 0970-1249.

[39] M. Hindry and N. Ratazzi. Points de torsion sur les variétés abéliennes de type GSp. *J. Inst. Math. Jussieu*, 11(1):27–65, 2012. ISSN 1474-7480. doi: 10.1017/S147474801000023X. URL http://dx.doi.org/10.1017/S147474801000023X.

[40] M. Hindry and N. Ratazzi. Torsion pour les variétés abéliennes de type I et II. *ArXiv e-prints*, May 2015.

[41] J. W. P. Hirschfeld. *Finite projective spaces of three dimensions*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1985. ISBN 0-19-853536-8. Oxford Science Publications.

[42] J. E. Humphreys. *Introduction to Lie algebras and representation theory*, volume 9 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1978. ISBN 0-387-90053-5. Second printing, revised.

[43] J. E. Humphreys. *Modular representations of finite groups of Lie type*, volume 326 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006. ISBN 978-0-521-67454-6; 0-521-67454-9.

[44] T. Ichikawa. Algebraic groups associated with abelian varieties. *Mathematische Annalen*, 289 (1):133–142, 1991. URL http://eudml.org/doc/164773.

[45] M. I. Jacobson. Variétés abéliennes de dimension deux ayant pour algèbre d'endomorphismes une algèbre de quaternions indéfinie (in Russian). *Uspekhi Mat. Nauk*, 29:185–186, 1974.

[46] J. C. Jantzen. Low-dimensional representations of reductive groups are semisimple. In *Algebraic groups and Lie groups*, volume 9 of *Austral. Math. Soc. Lect. Ser.*, pages 255–266. Cambridge Univ. Press, Cambridge, 1997.

[47] Z. Jelonek. On the effective Nullstellensatz. *Invent. Math.*, 162(1):1–17, 2005. ISSN 0020-9910. doi: 10.1007/s00222-004-0434-8. URL http://dx.doi.org/10.1007/s00222-004-0434-8.

[48] T. Kawamura. The effective surjectivity of mod l Galois representations of 1- and 2-dimensional abelian varieties with trivial endomorphism ring. *Comment. Math. Helv.*, 78 (3):486–493, 2003. ISSN 0010-2571. doi: 10.1007/s00014-003-0768-7. URL http://dx.doi.org/10.1007/s00014-003-0768-7.

[49] P. Kleidman and M. Liebeck. *The subgroup structure of the finite classical groups*, volume 129 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1990. ISBN 0-521-35949-X. doi: 10.1017/CBO9780511629235. URL http://dx.doi.org/10.1017/CBO9780511629235.

[50] J. Klüners and G. Malle. A database for field extensions of the rationals. *LMS J. Comput. Math.*, 4:182–196 (electronic), 2001. ISSN 1461-1570. doi: 10.1112/S1461157000000851. URL http://dx.doi.org/10.1112/S1461157000000851.

[51] J. Kollár. Sharp effective Nullstellensatz. *J. Amer. Math. Soc.*, 1(4):963–975, 1988. ISSN 0894-0347. doi: 10.2307/1990996. URL http://dx.doi.org/10.2307/1990996.

[52] T. Kubota. On the field extension by complex multiplication. *Trans. Amer. Math. Soc.*, 118: 113–122, 1965. ISSN 0002-9947.

[53] J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem. In *Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975)*, pages 409–464. Academic Press, London, 1977.

[54] S. Lang. Algebraic groups over finite fields. *Amer. J. Math.*, 78:555–563, 1956. ISSN 0002-9327.

[55] S. Lang. *Complex multiplication*, volume 255 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1983. ISBN 0-387-90786-6. doi: 10.1007/978-1-4612-5485-0. URL http://dx.doi.org/10.1007/978-1-4612-5485-0.

[56] M. Larsen and A. Lubotzky. Representation growth of linear groups. *J. Eur. Math. Soc. (JEMS)*, 10(2):351–390, 2008. ISSN 1435-9855. doi: 10.4171/JEMS/113. URL http://dx.doi.org/10.4171/JEMS/113.

[57] M. Larsen and R. Pink. Determining representations from invariant dimensions. *Invent. Math.*, 102(2):377–398, 1990. ISSN 0020-9910. doi: 10.1007/BF01233432. URL http://dx.doi.org/10.1007/BF01233432.

[58] M. Larsen and R. Pink. On ℓ-independence of algebraic monodromy groups in compatible systems of representations. *Invent. Math.*, 107(3):603–636, 1992. ISSN 0020-9910. doi: 10.1007/BF01231904. URL http://dx.doi.org/10.1007/BF01231904.

[59] M. Larsen and R. Pink. Abelian varieties, l-adic representations, and l-independence. *Math. Ann.*, 302(3):561–579, 1995. ISSN 0025-5831. doi: 10.1007/BF01444508. URL http://dx.doi.org/10.1007/BF01444508.

[60] M. J. Larsen and R. Pink. Finite subgroups of algebraic groups. *J. Amer. Math. Soc.*, 24 (4):1105–1158, 2011. ISSN 0894-0347. doi: 10.1090/S0894-0347-2011-00695-4. URL http://dx.doi.org/10.1090/S0894-0347-2011-00695-4.

[61] E. Larson and D. Vaintrob. Determinants of subquotients of Galois representations associated with abelian varieties. *J. Inst. Math. Jussieu*, 13(3):517–559, 2014. ISSN 1474-7480. doi: 10.1017/S1474748013000182. URL http://dx.doi.org/10.1017/S1474748013000182. With an appendix by B. Conrad.

[62] S. Le Fourn. Surjectivity of Galois representations associated with quadratic Q-curves. *ArXiv e-prints*, 1212.4713, Dec. 2012.

[63] M. W. Liebeck. On the orders of maximal subgroups of the finite classical groups. *Proc. London Math. Soc. (3)*, 50(3):426–446, 1985. ISSN 0024-6115. doi: 10.1112/plms/s3-50.3.426. URL http://dx.doi.org/10.1112/plms/s3-50.3.426.

[64] Q. Liu. *Algebraic geometry and arithmetic curves*, volume 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002. ISBN 0-19-850284-2. Translated from the French by Reinie Erné, Oxford Science Publications.

[65] F. Lübeck. Small degree representations of finite Chevalley groups in defining characteristic. *LMS J. Comput. Math.*, 4:135–169 (electronic), 2001. ISSN 1461-1570. doi: 10.1112/S1461157000000838. URL http://dx.doi.org/10.1112/S1461157000000838.

[66] L. Mai. Lower bounds for the ranks of CM types. *Journal of Number Theory*, 32(2):192 – 202, 1989. ISSN 0022-314X. doi: http://dx.doi.org/10.1016/0022-314X(89)90025-5. URL http://www.sciencedirect.com/science/article/pii/0022314X89900255.

[67] D. W. Masser. Counting points of small height on elliptic curves. *Bull. Soc. Math. France*, 117(2):247–265, 1989. ISSN 0037-9484. URL http://www.numdam.org/item?id=BSMF_1989__117_2_247_0.

[68] D. W. Masser. Multiplicative isogeny estimates. *J. Austral. Math. Soc. Ser. A*, 64(2):178–194, 1998. ISSN 0263-6115.

[69] D. W. Masser and G. Wüstholz. Some effective estimates for elliptic curves. In *Arithmetic of complex manifolds (Erlangen, 1988)*, volume 1399 of *Lecture Notes in Math.*, pages 103–109. Springer, Berlin, 1989. doi: 10.1007/BFb0095971. URL http://dx.doi.org/10.1007/BFb0095971.

[70] D. W. Masser and G. Wüstholz. Periods and minimal abelian subvarieties. *Ann. of Math. (2)*, 137(2):407–458, 1993. ISSN 0003-486X. doi: 10.2307/2946542. URL http://dx.doi.org/10.2307/2946542.

[71] D. W. Masser and G. Wüstholz. Galois properties of division fields of elliptic curves. *Bull. London Math. Soc.*, 25(3):247–254, 1993. ISSN 0024-6093. doi: 10.1112/blms/25.3.247. URL http://dx.doi.org/10.1112/blms/25.3.247.

[72] D. W. Masser and G. Wüstholz. Isogeny estimates for abelian varieties, and finiteness theorems. *Ann. of Math. (2)*, 137(3):459–472, 1993. ISSN 0003-486X. doi: 10.2307/2946529. URL http://dx.doi.org/10.2307/2946529.

[73] D. W. Masser and G. Wüstholz. Refinements of the Tate conjecture for abelian varieties. In *Abelian varieties (Egloffstein, 1993)*, pages 211–223. de Gruyter, Berlin, 1995.

[74] D. W. Masser and U. Zannier. Torsion anomalous points and families of elliptic curves. *Amer. J. Math.*, 132(6):1677–1691, 2010. ISSN 0002-9327.

[75] M. Matignon. Vers un algorithme pour la réduction stable des revêtements p-cycliques de la droite projective sur un corps p-adique. *Math. Ann.*, 325(2):323–354, 2003. ISSN 0025-5831. doi: 10.1007/s00208-002-0387-4. URL http://dx.doi.org/10.1007/s00208-002-0387-4.

[76] G. J. McNinch. Dimensional criteria for semisimplicity of representations. *Proc. London Math. Soc. (3)*, 76(1):95–149, 1998. ISSN 0024-6115. doi: 10.1112/S0024611598000045. URL http://dx.doi.org/10.1112/S0024611598000045.

[77] J. S. Milne. Basic theory of affine group schemes, 2012. Available at www.jmilne.org/math/.

[78] H. H. Mitchell. The subgroups of the quaternary abelian linear group. *Trans. Amer. Math. Soc.*, 15(4):379–396, 1914. ISSN 0002-9947. doi: 10.2307/1988794. URL http://dx.doi.org/10.2307/1988794.

[79] B. J. J. Moonen. Notes on Mumford-Tate Groups. Available at http://www.science.uva.nl/~bmoonen/NotesMT.pdf.

[80] B. J. J. Moonen and Y. G. Zarhin. Hodge and Tate classes on simple Abelian fourfolds. *Duke Math. J*, 77:553–581, 1995.

[81] B. J. J. Moonen and Y. G. Zarhin. Hodge classes on abelian varieties of low dimension. *Math. Ann.*, 315(4):711–733, 1999. ISSN 0025-5831. doi: 10.1007/s002080050333. URL http://dx.doi.org/10.1007/s002080050333.

[82] B. J. J. Moonen and Y. G. Zarhin. Hodge classes on abelian varieties of low dimension. *Mathematische Annalen*, 315:711–733, 1999. doi: 10.1007/s002080050333.

[83] L. Moret-Bailly. Pinceaux de variétés abéliennes. *Astérisque*, (129):266, 1985. ISSN 0303-1179.

[84] D. Mumford. Families of abelian varieties. In *Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965)*, pages 347–351. Amer. Math. Soc., Providence, R.I., 1966.

[85] D. Mumford. A note of Shimura's paper "Discontinuous groups and abelian varieties". *Math. Ann.*, 181:345–351, 1969. ISSN 0025-5831.

[86] D. Mumford. Abelian varieties. In *Studies in Mathematics, No. 5, Published for the Tata Institute of Fundamental Research*, 1970.

[87] V. K. Murty. Exceptional hodge classes on certain abelian varieties. *Mathematische Annalen*, 268(2):197–206, 1984. ISSN 0025-5831. doi: 10.1007/BF01456085. URL http://dx.doi.org/10.1007/BF01456085.

[88] E. Nart and X. Xarles. Additive reduction of algebraic tori. *Arch. Math. (Basel)*, 57(5): 460–466, 1991. ISSN 0003-889X. doi: 10.1007/BF01246743. URL http://dx.doi.org/10.1007/BF01246743.

[89] J. Neukirch. *Class field theory*, volume 280 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1986. ISBN 3-540-15251-2. doi: 10.1007/978-3-642-82465-4. URL http://dx.doi.org/10.1007/978-3-642-82465-4.

[90] J. Neukirch, A. Schmidt, and K. Wingberg. *Cohomology of Number Fields*. Grundlehren der mathematischen Wissenschaften. Springer, 2013. ISBN 9783540378891. URL http://books.google.fr/books?id=ZX8CAQAAQBAJ.

[91] R. Noot. Abelian varieties—Galois representation and properties of ordinary reduction. *Compositio Math.*, 97(1-2):161–171, 1995. ISSN 0010-437X. URL http://www.numdam.org/item?id=CM_1995__97_1-2_161_0. Special issue in honour of Frans Oort.

[92] J. Œsterlé. Versions effectives du théorème de Chebotarev sous l'hypothèse de Riemann généralisée. In *Journées Arithmétiques de Luminy (Colloq. Internat. CNRS, Centre Univ. Luminy, Luminy, 1978)*, volume 61 of *Astérisque*, pages 165–167. Soc. Math. France, Paris, 1979.

[93] M. Ohta. On l-adic representations of Galois groups obtained from certain two-dimensional abelian varieties. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 21:299–308, 1974. ISSN 0040-8980.

[94] T. Ono. Arithmetic of algebraic tori. *Ann. of Math. (2)*, 74:101–139, 1961. ISSN 0003-486X.

[95] F. Pazuki. Décompositions en hauteurs locales. *ArXiv e-prints*, May 2012.

[96] F. Pazuki. Heights, ranks and regulators of abelian varieties. *ArXiv e-prints*, June 2015.

[97] R. Pink. Classification of pro-p subgroups of SL_2 over a p-adic ring, where p is an odd prime. *Compositio Math.*, 88(3):251–264, 1993. ISSN 0010-437X. URL http://www.numdam.org/item?id=CM_1993__88_3_251_0.

[98] R. Pink. ℓ-adic algebraic monodromy groups, cocharacters, and the Mumford-Tate conjecture. *J. reine angew. Math*, (495), 1998.

[99] R. Pink. On Weil restriction of reductive groups and a theorem of Prasad. *Mathematische Zeitschrift*, 248(3):449–457, 2004. ISSN 0025-5874. doi: 10.1007/s002090100339. URL http://dx.doi.org/10.1007/s002090100339.

[100] I. I. Pjateckiĭ-Šapiro. Interrelations between the Tate and Hodge hypotheses for abelian varieties. *Mat. Sb. (N.S.)*, 85(127):610–620, 1971.

[101] H. Pohlmann. Algebraic cycles on abelian varieties of complex multiplication type. *Ann. of Math. (2)*, 88:161–180, 1968. ISSN 0003-486X.

[102] A. A. Premet. Weights of infinitesimally irreducible representations of Chevalley groups over a field of prime characteristic. *Mat. Sb. (N.S.)*, 133(175)(2):167–183, 271, 1987. ISSN 0368-8666.

[103] J. J. Ramón Marí. On the Hodge conjecture for products of certain surfaces. *Collect. Math.*, 59(1):1–26, 2008. ISSN 0010-0757. doi: 10.1007/BF03191179. URL http://dx.doi.org/10.1007/BF03191179.

[104] M. Raynaud. Schémas en groupes de type $(p, \dots, p)$. *Bull. Soc. Math. France*, 102:241–280, 1974. ISSN 0037-9484.

[105] M. Raynaud. Hauteurs et isogénies. *Astérisque*, (127):199–234, 1985. ISSN 0303-1179. Séminaire sur les pinceaux arithmétiques (Paris, 1983/84).

[106] M. Raynaud. p-groupes et réduction semi-stable des courbes. In *The Grothendieck Festschrift, Vol. III*, volume 88 of *Progr. Math.*, pages 179–197. Birkhäuser Boston, Boston, MA, 1990. doi: 10.1007/978-0-8176-4576-2_7. URL http://dx.doi.org/10.1007/978-0-8176-4576-2_7.

[107] I. Reiner. *Maximal orders*, volume 28 of *London Mathematical Society Monographs. New Series*. The Clarendon Press Oxford University Press, Oxford, 2003. ISBN 0-19-852673-3. Corrected reprint of the 1975 original, With a foreword by M. J. Taylor.

[108] K. A. Ribet. On ℓ-adic representations attached to modular forms. *Invent. Math.*, 28:245–275, 1975. ISSN 0020-9910.

[109] K. A. Ribet. Galois action on division points of Abelian varieties with real multiplications. *Amer. J. Math.*, 98(3):751–804, 1976. ISSN 0002-9327.

[110] K. A. Ribet. Division fields of abelian varieties with complex multiplication. *Mém. Soc. Math. France (N.S.)*, (2):75–94, 1980/81. ISSN 0583-8665. Abelian functions and transcendental numbers (Colloq., École Polytech., Palaiseau, 1979).

[111] K. A. Ribet. Generalization of a theorem of Tankeev. In *Seminar on Number Theory, 1981/1982*, pages Exp. No. 17, 4. Univ. Bordeaux I, Talence, 1982.

[112] K. A. Ribet. Hodge classes on certain types of Abelian varieties. *Amer. J. Math*, 105:523–538, 1983.

[113] K. A. Ribet. Images of semistable Galois representations. *Pacific J. Math.*, Special Issue: 277–297, 1997. ISSN 0030-8730. doi: 10.2140/pjm.1997.181.277. URL http://dx.doi.org/10.2140/pjm.1997.181.277. Olga Taussky-Todd: in memoriam.

[114] S. Sen. Lie algebras of Galois groups arising from Hodge-Tate modules. *Ann. of Math. (2)*, 97:160–170, 1973. ISSN 0003-486X.

[115] J.-P. Serre. *Corps locaux*. Hermann, Paris, 1968. Deuxième édition, Publications de l'Université de Nancago, No. VIII.

[116] J.-P. Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15(4):259–331, 1972. ISSN 0020-9910.

[117] J.-P. Serre. Quelques applications du théorème de densité de Chebotarev. *Inst. Hautes Études Sci. Publ. Math.*, (54):323–401, 1981. ISSN 0073-8301. URL http://www.numdam.org/item?id=PMIHES_1981__54__323_0.

[118] J.-P. Serre. Résumé des cours de 1985-1986, Annuaire du Collge de France, 1986.

[119] J.-P. Serre. *Abelian ℓ-adic Representations and Elliptic Curves*. A. K. Peters Ltd, 3 edition, Oct. 1997. ISBN 1568810776. URL http://www.worldcat.org/isbn/1568810776.

[120] J.-P. Serre. Letter to M-F. Vigneras, January 1^{st}, 1983. In *Œuvres. Collected papers. IV*. Springer-Verlag, Berlin, 2000.

[121] J.-P. Serre. Résumé des cours au Collège de France, 1984-1985. In *Œuvres. Collected papers. IV*, pages viii+657. Springer-Verlag, Berlin, 2000. 1985–1998.

[122] J.-P. Serre. Letter to K. Ribet, March 7^{th}, 1986. In *Œuvres. Collected papers. IV*. Springer-Verlag, Berlin, 2000.

[123] J.-P. Serre. Un critère d'indépendance pour une famille de représentations ℓ-adiques. *Comment. Math. Helv.*, 88(3):541–554, 2013. ISSN 0010-2571. doi: 10.4171/CMH/295. URL http://dx.doi.org/10.4171/CMH/295.

[124] J.-P. Serre and J. Tate. Good reduction of abelian varieties. *Ann. of Math. (2)*, 88:492–517, 1968. ISSN 0003-486X.

[125] S. S. Shatz. *Profinite groups, arithmetic, and geometry*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1972. Annals of Mathematics Studies, No. 67.

[126] G. Shimura and Y. Taniyama. *Complex multiplication of abelian varieties and its applications to number theory*, volume 6 of *Publications of the Mathematical Society of Japan*. The Mathematical Society of Japan, Tokyo, 1961.

[127] T. Shioda. Algebraic cycles on abelian varieties of Fermat type. *Math. Ann*, (258), 1981.

[128] A. Silverberg. Torsion points on abelian varieties of CM-type. *Compositio Math.*, 68(3):241–249, 1988. ISSN 0010-437X. URL http://www.numdam.org/item?id=CM_1988__68_3_241_0.

[129] A. Silverberg. Fields of definition for homomorphisms of abelian varieties. *J. Pure Appl. Algebra*, 77(3):253–262, 1992. ISSN 0022-4049. doi: 10.1016/0022-4049(92)90141-2. URL http://dx.doi.org/10.1016/0022-4049(92)90141-2.

[130] R. Steinberg. Representations of algebraic groups. *Nagoya Math. J.*, 22:33–56, 1963. ISSN 0027-7630.

[131] R. Steinberg. *Lectures on Chevalley groups*. Yale University, New Haven, Conn., 1968. Notes prepared by John Faulkner and Robert Wilson.

[132] R. Taylor and A. Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141(3):553–572, 1995. ISSN 0003-486X. doi: 10.2307/2118560. URL http://dx.doi.org/10.2307/2118560.

[133] D. Testerman and A. Zalesski. Subgroups of simple algebraic groups containing regular tori, and irreducible representations with multiplicity 1 non-zero weights. *ArXiv e-prints*, Mar. 2014.

[134] J. Tsimerman. Brauer-Siegel for arithmetic tori and lower bounds for Galois orbits of special points. *J. Amer. Math. Soc.*, 25(4):1091–1117, 2012. ISSN 0894-0347. doi: 10.1090/S0894-0347-2012-00739-5. URL http://dx.doi.org/10.1090/S0894-0347-2012-00739-5.

[135] E. Ullmo and A. Yafaev. Mumford-Tate and generalised Shafarevich conjectures. *Annales mathématiques du Québec*, 37(2):255–284, 2013. ISSN 2195-4755. doi: 10.1007/s40316-013-0009-4. URL http://dx.doi.org/10.1007/s40316-013-0009-4.

[136] A. Vasiu. Surjectivity criteria for p-adic representations. I. *Manuscripta Math.*, 112(3):325–355, 2003. ISSN 0025-2611. doi: 10.1007/s00229-003-0402-4. URL http://dx.doi.org/10.1007/s00229-003-0402-4.

[137] A. Vasiu. Some cases of the Mumford-Tate conjecture and Shimura varieties. *Indiana Univ. Math. J.*, 57(1):1–75, 2008. ISSN 0022-2518. doi: 10.1512/iumj.2008.57.3513. URL http://dx.doi.org/10.1512/iumj.2008.57.3513.

[138] V. E. Voskresenskiĭ. *Algebraic groups and their birational invariants*, volume 179 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1998. ISBN 0-8218-0905-9. Translated from the Russian manuscript by Boris Kunyavski [Boris È. Kunyavskiĭ].

[139] N. R. Wallach. On a theorem of Milnor and Thom. In *Topics in geometry*, volume 20 of *Progr. Nonlinear Differential Equations Appl.*, pages 331–348. Birkhäuser Boston, Boston, MA, 1996.

[140] W. C. Waterhouse. *Introduction to affine group schemes*, volume 66 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1979. ISBN 0-387-90421-2.

[141] B. S. Webb, editor. *Surveys in combinatorics 2005*, volume 327 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2005. ISBN 978-0-521-61523-2; 0-521-61523-2. doi: 10.1017/CBO9780511734885. URL http://dx.doi.org/10.1017/CBO9780511734885. Papers from the 20th British Combinatorial Conference held at the University of Durham, Durham, July 10–15, 2005.

[142] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)*, 141(3):443–551, 1995. ISSN 0003-486X. doi: 10.2307/2118559. URL http://dx.doi.org/10.2307/2118559.

[143] J. Wilson. Curves of genus 2 with real multiplication by a square root of 5. *PhD Thesis*, 1998. Available at http://eprints.maths.ox.ac.uk/32/.

[144] B. Winckler. Letter to the author.

[145] J.-P. Wintenberger. Démonstration d'une conjecture de Lang dans des cas particuliers. *J. Reine Angew. Math.*, 553:1–16, 2002. ISSN 0075-4102. doi: 10.1515/crll.2002.099. URL http://dx.doi.org/10.1515/crll.2002.099.

[146] S. Wong. Twists of Galois representations and projective automorphisms. *J. Number Theory*, 74(1):1–18, 1999. ISSN 0022-314X. doi: 10.1006/jnth.1998.2307. URL http://dx.doi.org/10.1006/jnth.1998.2307.

[147] W. J. Wong. Irreducible modular representations of finite Chevalley groups. *J. Algebra*, 20: 355–367, 1972. ISSN 0021-8693.

[148] Y. G. Zarhin. Endomorphisms of Abelian varieties over fields of finite characteristic. *Izv. Akad. Nauk SSSR Ser. Mat.*, 39(2):272–277, 471, 1975. ISSN 0373-2436.

[149] Y. G. Zarhin. Abelian varieties in characteristic p. *Mat. Zametki*, 19(3):393–400, 1976. ISSN 0025-567X.

[150] Y. G. Zarhin. Torsion of abelian varieties in finite characteristic. *Math. Notes*, 22(1):493–498, 1978. ISSN 0025-567X.

[151] Y. G. Zarhin. Abelian varieties, l-adic representations and Lie algebras. Rank independence on l. *Invent. Math.*, 55(2):165–176, 1979. ISSN 0020-9910. doi: 10.1007/BF01390088. URL http://dx.doi.org/10.1007/BF01390088.

[152] Y. G. Zarhin. Weights of simple Lie algebras in the cohomology of algebraic varieties. *Izv. Akad. Nauk SSSR Ser. Mat.*, 48(2):264–304, 1984. ISSN 0373-2436.

[153] Y. G. Zarhin. Hyperelliptic Jacobians without complex multiplication. *Math. Res. Lett.*, 7(1): 123–132, 2000. ISSN 1073-2780. doi: 10.4310/MRL.2000.v7.n1.a11. URL http://dx.doi.org/10.4310/MRL.2000.v7.n1.a11.

[154] Y. G. Zarhin. Abelian varieties without homotheties. *Math. Res. Lett.*, 14(1):157–164, 2007. ISSN 1073-2780. doi: 10.4310/MRL.2007.v14.n1.a13. URL http://dx.doi.org/10.4310/MRL.2007.v14.n1.a13.

[155] B. Zhao. On the Mumford-Tate conjecture of abelian fourfolds, 2013. Available at http://www.math.ucla.edu/~zhaobin/MTAV4.pdf.

[156] D. Zywina. Bounds for Serre's open image theorem. *ArXiv e-prints*, 1102.4656, Feb. 2011.

[157] D. Zywina. An explicit Jacobian of dimension 3 with maximal Galois action. 2015. Available from http://www.math.cornell.edu/~zywina/.