BOUNDS FOR SERRE'S OPEN IMAGE THEOREM FOR ELLIPTIC CURVES OVER NUMBER FIELDS

DAVIDE LOMBARDO

Département de Mathématiques d'Orsay

Abstract

For an elliptic curve E/K without potential complex multiplication we bound the index of the image of $\mathrm{Gal}\left(\overline{K}/K\right)$ in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$, the representation being given by the action on the Tate modules of E at the various primes. The bound is explicit and only depends on $[K : \mathbb{Q}]$ and on the stable Faltings height of E. We also prove a result relating the structure of closed subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ to certain Lie algebras naturally attached to them.

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 $\mathrm{Gal}\left(\overline{K}/K\right)$ 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's and whose image we denote G_∞. By a theorem of Serre ([Ser72, §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 [Zyw11] in the special case $K = \mathbb{Q}$. More precisely we show:

Theorem 1.1. *Let E/K be an elliptic curve that does not admit complex multiplication. The inequality*

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

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

E-mail address: `davide.lombardo@math.u-psud.fr`.

Remark 1.2. We actually prove a more precise result (theorem 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 2.1; unpublished results that Eric Gaudron and Gaël Rémond have been kind enough to share with the author show that the statement still holds with the much better constants $C_1 = \exp\left(6.4 \cdot 10^{13}\right)$ and $C_2 = 24480$, cf. remark 9.4.

As an easy corollary we also get:

Corollary 1.3. *Let E/K be an elliptic curve that does not admit complex multiplication. There exists a constant $C(E/K)$ with the following property: for every $x \in E_{\mathrm{tors}}(\overline{K})$ (of order denoted $N(x)$) the inequality*

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

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

Remark 1.4. This corollary (with the same proof, but with a non-effective $C(E/K)$) follows directly from the aforementioned theorem of Serre ([Ser72, §4, Théorème 3]). The exponent 2 for $N(x)$ is best possible, as it 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 ([Mas89, p. 262]) proves an inequality of the form

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

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

The route we follow lends itself to a comparison with [Zyw11], so let us review very briefly the argument of this paper. By a technique due to Masser and Wüstholz (cf. [MW93a], [MW93b] and [Mas98]), 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 rids us of 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), and leaves us to deal with the case of non-surjective reduction, modulo a finite number of 'small' primes.

In [Zyw11] 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 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 p-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 [Zyw11], 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 paper reflects the steps alluded to above: in section 2 we recall an explicit form of the isogeny theorem (as proved by Gaudron and Rémond in [GR14] 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 4 (for odd ℓ) and 5 (for $\ell = 2$):

Theorem 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 $\left\{ g - \frac{\mathrm{tr}(g)}{2} \cdot \mathrm{Id} \mid g \in G \right\}$.*

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 120 (resp. with trivial reduction modulo 2 and of index at most 192 for $\ell = 2$), with the following property: if $L(H_1)$ contains $\ell^s \mathfrak{sl}_2(\mathbb{Z}_\ell)$ for a certain integer $s \geq 1$, then H_1 itself contains

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

The methods of these sections are then applied in section 7 to get bounds valid for *every* prime ℓ (cf. theorem 7.5, which might have some independent interest), while section 8 deals with the large primes through the aforementioned ideas of Masser and Wüstholz. Finally, in section 9 we put it all together to get the adelic estimate.

2. PRELIMINARIES ON ISOGENY BOUNDS

The main tool that makes all the effective estimates possible is a very explicit isogeny-type theorem taken from [GR14], which builds on the seminal work of Masser and Wüstholz (cf. [MW93c] and [MW93b]). 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 2.1. *([GR14] 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 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 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 $10^{13N}[K:\mathbb{Q}]^{2N} \max\left(h(E), \log[K:\mathbb{Q}], 1 \right)^{2N}$.

In [Mas98] Masser shows the following (cf. especially lemma 3.4 of *op. cit.*):

Theorem 2.4. *(Masser) Suppose that A/K is an Abelian variety that is isomorphic over K to a product $A_1^{e_1} \times \ldots \times A_m^{e_m}$, where each A_i is simple and has trivial endomorphism ring over 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 an isogeny $A^* \to A$ of degree dividing b_0.*

We will denote $b_0(K, A)$ 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 through all the finite

extensions of K of degree bounded by d. On one hand, $b_0(K, A)$ divides $b_0(K', A)$ ([Mas98, p.190]); on the other $b_0(K', A) \leq b(d[K : \mathbb{Q}], h(A), \dim(A))$ stays bounded, and therefore the number

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

exists and is finite. The function $b_0(K, A; d)$ is studied in [Mas98], Theorem D, mostly through the following elementary lemma:

Lemma 2.5. *([Mas98, Lemma 7.1]) Let $B, C \geq 1$ 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 $\operatorname{lcm}(A) \leq BC^t$. The least common multiple of the elements of $\mathcal{B}$ is then finite, and does not exceed $4^{eC}B^{1+\log(C)}$, where $e = \exp(1)$.*

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 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 $L = K_1 \cdots K_t$. Observe 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 as the right hand side is clearly at least 1, it is enough to show that

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

and since $\log(d) > 0$ we have $(1 + \log(d))^t \geq 1 + t\log(d)$ by Bernoulli's inequality, and the claim follows. Therefore $\operatorname{lcm}(b_0(K_1, A), \ldots, b_0(K_t, A))$ 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

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

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

Remark 2.7. We are only going to use the function $b_0(K, A; d)$ for bounded values of d (in fact, $d \leq 120$), 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 2.1 mentioned in remark 2.3 we get

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

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

Theorem 2.8. *Suppose A/K is an Abelian variety isomorphic over K to a product of simple 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 we can choose 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 finite and is bounded by a polynomial in $b([K : \mathbb{Q}], \dim(A), h(A))$.

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 from Pink's paper [Pin93], 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.

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

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

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

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

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

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 $x R_b \in \mathcal{B}_\ell(n)$ is diagonal, and since every diagonal matrix in $\mathcal{B}_\ell(n)$ is by definition of the form C_c for a certain $c \in \ell^n \mathbb{Z}_\ell$ we are done. $\qquad\square$

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 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 congruences $\sqrt{1+x} \equiv 1 \pmod 4$ (resp. modulo ℓ) now follow.   □

Lemma 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 3.1) we now just need to show that $\mathcal{B}_\ell(n)'$ contains C_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 $MNM^{-1}N^{-1} = \begin{pmatrix} 1+c & 0 \\ 0 & \frac{1}{1+c} \end{pmatrix} = C_c$. 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 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* $C_c = \begin{pmatrix} 1 + c & 0 \\ 0 & \frac{1}{1+c} \end{pmatrix}$, *then* G

contains C_{cu} *for every* $u \in \mathbb{Z}_2$.

In particular, 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* C_c, *then* G *contains* $\mathcal{B}_2(s)$.

Proof. We show that the set W made of the w's 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 C_{cu} is slightly different, in that $u \mapsto C_{cu}$ is not a group morphism; however, if $w \in \mathbb{Z}_2$, then we see that

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

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 C_γ belongs to G. The last statement is now an immediate consequence of lemma 3.1. $\square$

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 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 $\mathfrak{sl}_2(A) = \{ M \in M_2(A) \mid \mathrm{tr}(M) = 0 \}$ 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 [Pin93]:

Definition 3.6. *Let* G *be a 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*

$$\Theta : \quad G \quad \to \quad \mathfrak{sl}_2(\mathbb{Z}_\ell)$$
$$g \quad \mapsto \quad g - \tfrac{1}{2} \mathrm{tr}(g) \cdot \mathrm{Id}.$$

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 3.7. *The Lie algebra associated with* 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* $\mathbb{Z}_\ell$ *topoogically generated by all the traces* $\mathrm{tr}(xy)$ *for* x, y *in* $L(G)$.

Remark 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 3.9. *([Pin93, 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*

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

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 3.10. *([Pin93, Formula 1.3]) For every $g_1, g_2 \in \mathrm{GL}_2(\mathbb{Z}_\ell)$, if $\ell \neq 2$ (resp. for every $g_1, g_2 \in \left\{ x \in \mathrm{GL}_2(\mathbb{Z}_2) \mid \mathrm{tr}(x) \equiv 0 \pmod 2 \right\}$, 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 $G(\ell^n)$ 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 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 π_k the projection $A^n \to A$ on the k-th component. There exist a basis $x_1, \ldots, x_m$ 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's 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 v be the discrete valuation associated with A; the set $\left\{ v(\pi_1(x)) \mid x \in M \right\}$ consists of non-negative integers, therefore it admits a minimum k_1. Take x_1 to be any element of M such that $v(\pi_1(x_1)) = k_1$. 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 $\pi_1 \circ f(m) \equiv 0$. 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 $f(x_2), \ldots, f(x_m)$ 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 think the u_i's as elements of A^{n+1}. It is now enough to show that, with this choice of the x_i's, 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 $\sigma_{i+1,1} = \dfrac{\pi_1(x_{i+1})}{\pi_1(x_1)}, \sigma_{ij} = \tau_{ij}$.

As for the last statement, observe that the matrix giving the transformation from the x_i's to the t_j's is unitriangular, hence invertible. $\qquad\square$

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 3.12. *A subgroup L of $\mathrm{GL}_2(\mathbb{F}_\ell)$ is said to be:*

- ***split Cartan**, if L is conjugated to the subgroup of diagonal matrices. In this case the order of L 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 $L = A^\times$. The order of L is prime to ℓ, and L 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 L is the normalizer of $\mathcal{C}$. The index $[L : \mathcal{C}]$ is 2, and ℓ does not divide the order of L.*
- ***Borel**, if L is conjugated to the subgroup of upper-triangular matrices. In this case L has a unique ℓ-Sylow, consisting of the matrices of the form $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$.*
- ***exceptional**, if the image $\mathbb{P}L$ of L in $\mathbb{P}\mathrm{GL}_2(\mathbb{F}_\ell)$ is isomorphic to either A_4, S_4 or S_5, in which case the order of $\mathbb{P}L$ 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 3.13. *(Dickson's classification, cf. [Ser72]) Let ℓ be a prime number and L be a subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$. Then we have:*

- *if ℓ divides the order of L, then either L contains $\mathrm{SL}_2(\mathbb{F}_\ell)$ or it is contained in a Borel subgroup;*
- *if ℓ does not divide the order of L, then L 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).

Convention. Let L be a subgroup of $\mathrm{GL}_2(\mathbb{F}_\ell)$. In order to simplify the wording it is useful to take the following conventions: we will say that L is contained in a Borel subgroup only when this is true and in addition the order of L is divisible by ℓ, and similarly we will say that L is contained in the normalizer of a Cartan subgroup only when this is true and at the same time L is *not* contained in any Cartan subgroup.

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 3.14. *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 $\mathrm{SL}_2(\mathbb{F}_\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$: 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 [Ser97, 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. $\square$

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 3.15. *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 $G^{\det=1}$ the group $G \cap \mathrm{SL}_2(\mathbb{Z}_\ell)$ (resp. $G \cap \mathrm{SL}_2(\mathbb{F}_\ell)$).*

Lemma 3.16. *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 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 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 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 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. $\qquad\square$

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 3.17. *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 $G(\ell) \cong \frac{G}{N(G)}$.*

(2) *If $G(\ell)$ 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 $G(\ell) \cong \frac{G}{N(G)}$.*

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

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

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

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 4.1. *Let ℓ be an odd prime, s a natural number and G a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_\ell)$.*

(i) *Suppose that $G(\ell)$ is contained in a Cartan or Borel subgroup, and that $|G/N(G)| \neq 4$. Then G has the following property:*
$$(\star) \quad \textit{if } L(G) \textit{ contains } \ell^s \mathfrak{sl}_2(\mathbb{Z}_\ell), \textit{ then } L(N(G)) \textit{ contains } \ell^{2s}\mathfrak{sl}_2(\mathbb{Z}_\ell).$$

(ii) *Without any assumption on G, there is a closed subgroup H of G that satisfies $[G : H] \leq 60$ and the conditions in (i) (so H has property $(\star)$).*

Theorem 4.2. *Let ℓ be an odd prime, s a natural number and G a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$.*

(i) *Suppose that G satisfies the two conditions:*
  (a) $\det(g)$ *is a square in $\mathbb{Z}_\ell$ for every $g \in G$;*
  (b) $\mathrm{Sat}(G)^{\det=1}$ *satisfies the hypotheses of theorem 4.1 (i).*
    Then G has the following property:
$$(\star\star) \quad \textit{if } L(G) \textit{ contains } \ell^s\mathfrak{sl}_2(\mathbb{Z}_\ell), \textit{ then } G' \textit{ contains } \mathcal{B}_\ell(4s).$$

(ii) *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 $[G : H] \leq 120$ and the conditions in (i) (so H has property $(\star\star)$).*

Remark 4.3. Let us make condition (b) in this theorem a little more explicit. By the description of the maximal normal pro-ℓ subgroup given in lemma 3.17, the conditions on G can be read on $(\mathrm{Sat}(G))^{\det=1}(\ell)$ as follows: $(\mathrm{Sat}(G))^{\det=1}(\ell)$ should either be a cyclic group or of Borel type; in the first case we ask that its order be different from 4, while in the second the condition reads $\left|\mathrm{Sat}(G)^{\det=1}(\ell)/S\right| \neq 4$, where S is the unique ℓ-Sylow of $\mathrm{Sat}(G)^{\det=1}(\ell)$.

Let us remark that the statements numbered (ii) in the above theorems require a case by case analysis, which will be carried out in section 4.6 for theorem 4.2 (the proof of theorem 4.1 (ii) is perfectly analogous). In the same section we will also show that part (i) of theorem 4.2 can be reduced to the corresponding statement in theorem 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 4.1 (i), 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 C the operator $x \mapsto g^{-1}xg$: then C 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 expect a decomposition $L(G) \cong \langle \Theta(g) \rangle \oplus M$ (M being a C-stable subspace of dimension 2), and the projection operator $p : L(G) \to M$ would be a polynomial in C. We would also expect M to consist of elements coming from $N(G)$, since the Lie algebra of $\langle g \rangle$ is simply $\langle \Theta(g) \rangle$. We would finally deduce the equality $L(N(G)) = \mathfrak{sl}_2(\mathbb{Q}_\ell)$ by exploiting the description of M as a (p, C)-module and the fact that $L(N(G))$ is both a Lie algebra and a C-module. 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 its Lie algebra not to be solvable, and therefore the same should hold for $L(G)$. But the only non-solvable subalgebra of $\mathfrak{sl}_2(\mathbb{Q}_\ell)$ is $\mathfrak{sl}_2(\mathbb{Q}_\ell)$ itself, so we expect $L(G)$ to be very large even if $N(G)$ is very small.

In what follows we prove (i) of theorem 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 4.6 we finish the proof of theorem 4.2.

4.1. **The case** $|G/N(G)| = 2$. Suppose first that $G(\ell)$ is Cartan, 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 $A \in M_2(\mathbb{Z}_\ell)$. 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 $\Theta(-g_i) = -\Theta(g_i)$, so $L(G) = L(N(G))$ and the claim follows.

Next suppose $G(\ell)$ is Borel. 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.

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 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's 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. $\qquad\square$

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

Lemma 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 algebra $L(N)$ is stable under C.*

Proof. As $\Theta(N)$ generates $L(N)$ it is enough to prove that C 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. $\qquad\square$

Lemma 4.6. *Let L be a C-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
$$C(l_1) = \begin{pmatrix} x_{11} & 0 \\ a^2 x_{21} & -x_{11} \end{pmatrix},$$
so L contains $\begin{pmatrix} x_{11} & 0 \\ a^2 x_{21} & -x_{11} \end{pmatrix} - l_1 = \begin{pmatrix} 0 & 0 \\ (a^2-1)x_{21} & 0 \end{pmatrix}$, 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
$$\left[\begin{pmatrix} 0 & (a^2-1)y_{12} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ (a^2-1)x_{21} & 0 \end{pmatrix}\right] = \begin{pmatrix} (a^2-1)^2 x_{21}y_{12} & 0 \\ 0 & -(a^2-1)^2 x_{21}y_{12} \end{pmatrix},$$
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
$$a^2 C(l_1) - l_1 = \begin{pmatrix} (a^2-1)x_{11} & 0 \\ (a^4-1)x_{21} & -(a^2-1)x_{11} \end{pmatrix}$$
and $a^{-2}C(l_2) - l_2$, and noticing that since $\ell \nmid a^4 - 1$ we have $v_\ell((a^4-1)x_{21}) = v_\ell(x_{21})$ and $v_\ell((a^{-4}-1)y_{12}) = v_\ell(y_{12})$). $\qquad\square$

We know from lemma 4.5 that $L(N(G))$ is C-stable, so in order to apply lemma 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 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.

4.3. **The Borel case.** Suppose $G(\ell)$ is included in a Borel subgroup (so, by our convention, its order is divisible by ℓ). 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.

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 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 $\mathbb{Z}_\ell \left[\sqrt{\varepsilon} \right]$. The two eigenvalues of g in $\mathbb{Z}_\ell \left[\sqrt{\varepsilon} \right]$ 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, we have that a, b map respectively to $[a], [b]$ modulo ℓ).

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 $\left(\mathbb{F}_\ell[\sqrt{[\varepsilon]}] \right)^2$ are therefore nonzero eigenvectors of $[g]$ corresponding to different eigenvalues, hence 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 $\begin{pmatrix} \mathbf{z} & | & \mathbf{w} \end{pmatrix}$ 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, against the assumptions. $\square$

We can also assume that G contains $-\mathrm{Id}$, since replacing G with $G \cdot \{\pm \mathrm{Id}\}$ does not alter neither the derived subgroup nor the Lie algebra of G. By lemma 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

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

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

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

Lemma 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. The Lie algebra $L(N(G))$ is C-stable by lemma 4.5, so the identity

$$(4.1) \qquad \frac{1}{2ab}\left(C \begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix} - (a^2 + b^2\varepsilon) \begin{pmatrix} -F & -\varepsilon E \\ E & F \end{pmatrix} \right) = \begin{pmatrix} -\varepsilon E & -\varepsilon F \\ F & \varepsilon E \end{pmatrix}$$

shows that $\begin{pmatrix} -\varepsilon E & -\varepsilon F \\ F & \varepsilon E \end{pmatrix}$ is in $L(N(G))$. At least one between 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

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

where φ, γ are not zero modulo ℓ. In this second case we have $\varepsilon E^2 - F^2 = \ell^{2v_\ell(E)}\left(\varepsilon\gamma^2 - \varphi^2\right)$, and $\left(\varepsilon\gamma^2 - \varphi^2\right)$ 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 (4.1) 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 (4.1) to this last matrix we finally deduce that $L(N(G))$ also contains $\begin{pmatrix} 0 & -\varepsilon F \\ F & 0 \end{pmatrix}$. $\qquad\square$

Lemma 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 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. $\qquad\square$

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 $g_2 g^k$ and $g^k g_2$ 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 $N(G)$, against 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 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.

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

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

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 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.

4.5. **Optimality.** The following examples show that it is neither possible to extend theorem 4.2 to the exceptional case nor to improve the exponent $2s$.

Proposition 4.10. *Let ℓ be a prime $\equiv 1 \pmod 4$. For every $t \in \mathbb{N}$ there exists a closed subgroup G of $\mathrm{SL}_2(\mathbb{Z}_\ell)$ whose Lie algebra is $\mathfrak{sl}_2(\mathbb{Z}_\ell)$ and whose maximal pro-ℓ subgroup is contained in $\mathcal{B}_\ell(t)$.*

Proof. The following six elements form a finite subgroup H of $\mathrm{PSL}_2(\mathbb{Z}[i])$ (isomorphic to S_3: it's the group of symmetries of $\{0, 1, \infty\} \subset \mathbb{P}^1(\mathbb{Z}[i])$):

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

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 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 [Pin93]

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

is a pro-ℓ group with 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 $g^{-1}Mg = M$, 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 $t \geq 2s$.

4.6. **Proof of theorem 4.2.** We now prove (i) of theorem 4.2 by reducing it to the corresponding statement in theorem 4.1.

As G and $\mathrm{Sat}(G)$ have the same 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 3.16 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 4.1 (i), 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 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 (i).

We are now left with the task of proving (ii). 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 60 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 3.13):

- if $\ell \geq 5$ and $G(\ell)$ contains $\mathrm{SL}_2(\mathbb{F}_\ell)$, then it follows from lemma 3.14 that $G' = \mathrm{SL}_2(\mathbb{Z}_\ell)$.
- if $\ell = 3$ we can consider the kernel H of $G \to \mathrm{GL}_2(\mathbb{F}_3)$. The index of H in G is at most

$$\left|\left\{g \in \mathrm{GL}_2(\mathbb{F}_3) \mid \det(g) \text{ is a square}\right\}\right| = 24,$$

and H satisfies the conditions in (i) since $H(\ell) = \{\mathrm{Id}\}$ (cf. remark 4.3).

- if $G(\ell)$ is exceptional we take $H = \ker(G \to G(\ell) \to \mathbb{P}G(\ell))$. The index of H in G is at most 60, and H satisfies the conditions in (i) since

$$\mathrm{Sat}(H)^{\det=1}(\ell) = \mathrm{Sat}(H(\ell))^{\det=1} = \{\pm \mathrm{Id}\}.$$

- if $G(\ell)$ is a (split or nonsplit) Cartan then the same is true for $\mathrm{Sat}(G)^{\det=1}(\ell)$. If $\left(\mathrm{Sat}(G)^{\det=1}\right)(\ell)$ does not have order 4 we are done, so suppose this is the case. Then $\mathbb{P}G(\ell)$ has at most 4 elements, and we can take

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

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 (i).

- if $G(\ell)$ is contained in the normalizer of a (split or nonsplit) Cartan subgroup $\mathcal{C}$, then H has a subgroup H_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 Borel then $\mathrm{Sat}(G)^{\det=1}(\ell)$ is again Borel; if we set $G_2 = \mathrm{Sat}(G)^{\det=1}$ the only additional condition we need to satisfy is $|G_2/N(G_2)| \neq 4$. If $|G_2/N(G_2)| = 4$ we consider the group morphism

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

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 $\varphi(\lambda g_2) = \varphi(g_2)$ we deduce $\varphi(G) = \varphi(G_2)$. On the other hand, when restricted to G_2 the function φ becomes

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

and we have already remarked that $g \mapsto [g] = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} \mapsto a$ is simply the quotient map $G_2 \twoheadrightarrow G_2/N(G_2)$. Hence φ factors through the quotient $G_2/N(G_2)$ and $|\varphi(G)| = |\varphi(G_2)| \leq 4$. We take H to be the kernel of φ. The inequality $[G:H] \leq 4$ is immediate, and the theorem applies to H. To check this last claim, note that $\ker \pi \subseteq H$, so $G/H \cong \frac{G/\ker \pi}{H/\ker \pi} = \frac{G(\ell)}{H(\ell)}$. Therefore $[G(\ell):H(\ell)]$ divides 4, which shows that $H(\ell)$ is still Borel, since its order is divisible by ℓ. On the other hand, 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.

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

   (i) Suppose that $G(4)$ is trivial and $\det(G) \equiv 1 \pmod 8$. Then G has the following property: if $L(G)$ contains $\ell^n \mathfrak{sl}_2(\mathbb{Z}_2)$ for a certain n, the derived subgroup G' of G contains the principal congruence subgroup $\mathcal{B}_2(12n + 2)$.

(ii) 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 (i).

Note that (ii) 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 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 5.2. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$ whose reduction modulo 4 is trivial. If its Lie algebra $L(G)$ contains $2^s \mathfrak{sl}_2(\mathbb{Z}_2)$ for a certain $s \geq 2$, then G itself 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 start with a compactness lemma. Our arguments will only ever 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 do belong to $\Theta(G)$.

Lemma 5.3. *Let $e \geq 2$ and $g \in G$. 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 $g = \dfrac{\mathrm{tr}(g)}{2} \mathrm{Id} + \Theta(g)$. Recall that G is a subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$, so 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 $\dfrac{\mathrm{tr}(g)}{2} = \sqrt{1 + a^2 + bc} = \sum_{j=0}^{\infty} \binom{1/2}{j} (a^2 + bc)^j$ by lemma 3.2. Given that $a^2 + bc \equiv 0 \pmod{2^{2e}}$ and $2e > 3$, using again lemma 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 $\mathcal{B}_2(2)$ through 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 $a^2 + bc = \frac{1}{2} \operatorname{tr}\left(\Theta(g)^2\right)$, so we can write

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

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. $\qquad\square$

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

Lemma 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 3.10 we find

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

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 $\operatorname{tr}(g_1) - 2 \equiv 0 \pmod{2^{2e_1}}$ and $\operatorname{tr}(g_2) - 2 \equiv 0 \pmod{2^{2e_2}}$ by lemma 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 $y = x_1 + x_2 + (\text{triangular error terms})$ is indeed triangular. $\qquad\square$

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

Lemma 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 $\lambda = \frac{\operatorname{tr}(g)}{2}$ is a unit in $\mathbb{Z}_2$. We can therefore form $\tilde{g} = \frac{1}{\lambda} g$, which certainly exists as a matrix in $\operatorname{GL}_2(\mathbb{Z}_2)$, even though it does not necessarily belong to G. Our choice of $\tilde{g}$ is made so as to ensure $\operatorname{tr}(\tilde{g}) = 2$, so the formula given in lemma 3.10 (applied with $g_1 = g_2 = \tilde{g}$) yields

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

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. $\qquad\square$

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 $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. The maps π_{ij} are obviously linear and continuous.

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 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 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 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 3.11. $\qquad\square$

As previously mentioned, in order to make the quadratic error terms appearing in lemma 5.4 negligible we need to work with matrices that are highly divisible by 2:

Lemma 5.8. *There exist $y_1, y_2, y_3 \in \Theta(G)$ and units $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{Z}_2^\times$ such that $y_i = \lambda_i \cdot 2^{4s} x_i$; 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 lemmas 5.5 and 5.6) 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's generate $L \supseteq 2^s\mathfrak{sl}_2(\mathbb{Z}_2)$, hence $M = 2^{4s}L$ contains $2^{5s}\mathfrak{sl}_2(\mathbb{Z}_2)$. $\qquad\square$

Notation. Let y_1, y_2, y_3 be given by the previous lemma. The properties of the x_i's become corresponding properties of the y_i's:

- 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 5.9. *The 2-adic valuations of a_{21}, b_{11} and c_{12} are at most $5s$.*

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 λ_i's in $\mathbb{Z}_2$. Comparing the bottom-left coefficient we have $\lambda_1 a_{21} = 2^{5s}$, 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. $\qquad\square$

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 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 addition 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 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 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 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

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

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 5.4. $\qquad\square$

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 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 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 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 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 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. $\qquad\square$

The result we were really aiming for follows at once:

Proposition 5.13. *Let G be a closed subgroup of $\mathrm{SL}_2(\mathbb{Z}_2)$ such that $L(G)$ contains $2^s\mathfrak{sl}_2(\mathbb{Z}_2)$ for a certain $s \geq 2$. 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 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 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 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 $\zeta \tilde{b}_{12} \equiv \dfrac{d_{11}}{b_{11}} \tilde{b}_{12}$ (mod 2^{6s-1}). Since furthermore $\tilde{b}_{12} \equiv b_{12}$ (mod 2^{6s-1}) we deduce $\zeta \tilde{b}_{12} \equiv \dfrac{d_{11}}{b_{11}} b_{12}$ (mod 2^{6s-1}). But then the inequality $v_2(c_{12}) \leq 5s$ (cf. remark 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 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 $2^s \tilde{b} \equiv 0$ (mod 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

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

$\square$

We are now ready for the proof of theorem 5.2:

Proof. With all the preliminaries in place this is now quite easy: by proposition 5.13 $\Theta(G)$ 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 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

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

is an operator of the form $C_c = \begin{pmatrix} 1+c & 0 \\ 0 & \frac{1}{c+1} \end{pmatrix}$, 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 $\{g^t \mid g \in G\}$ 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 C_c for certain a, b, c of valuation at most $6s$, so it follows from lemma 3.4 that G contains $\mathcal{B}_2(6s)$. $\square$

Remark 5.14. The above result should be thought of as an analogue of theorem 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 [Pin93] the problem seems to be substantially harder than for $\ell \neq 2$.

It is now easy to deduce theorem 5.1 (i):

Proof. The proof follows closely that of theorem 4.2 (i): we can replace G first by $H = G \cdot (1 + 8\mathbb{Z}_2)$ and then by $H_0 = G \cdot (1 + 8\mathbb{Z}_2) \cap \mathrm{SL}_2(\mathbb{Z}_2)$ 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 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 3.3. $\qquad\square$

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 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 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 $\min_{m \in L} v_\ell \circ \pi_{21}(m)$, where again π_{21} denotes the projection on the bottom-left coefficient in the standard matrix representation of elements of $\mathfrak{sl}_2(\mathbb{Z}_\ell)$. Furthermore, for every positive n, we let $L(\ell^n)$ will be the image of the mod-ℓ^n reduction map $\pi_n : L \to \mathfrak{sl}_2(\mathbb{Z}/\ell^n\mathbb{Z})$; clearly $L(\ell^n)$ is a Lie algebra over $\mathbb{Z}/\ell^n\mathbb{Z}$.

Remark 6.2. It is apparent from the very definition of a reduced basis that $k(L) = v_\ell(a_{21})$.

By definition, the images of x_1, x_2, x_3 in $L(\ell^n)$ generate it as a $(\mathbb{Z}/\ell^n\mathbb{Z})$-module. By a slight abuse of language, we say that $L(\ell^n)$ is generated by j elements, or that j is the number of generators of $L(\ell^n)$, if exactly j among the images of x_1, x_2, x_3 modulo ℓ^n are nonzero. Also recall that v denotes $v_\ell(2)$.

Proposition 6.3. *Suppose L as above is obtained as $\overline{\Theta(G)}$ for a certain pro-ℓ 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 $\mathrm{GL}_2(\mathbb{Z}/\ell^m\mathbb{Z})$. Fix an integer $n \geq 1$ and let j_n be the number of generators of $L(\ell^n)$. The following are the only possibilities:*
- *if j_n is at most 1 then $G(\ell^n)$ is abelian.*
- *if $j_n = 2$ then either $L(\ell^{2n})$ has 3 generators or $G(\ell^{n-k(L)+1-2v})$ is contained in the subgroup of upper-triangular matrices (up to a change of coordinates in $\mathrm{GL}_2(\mathbb{Z}_\ell)$).*
- *if $j_n = 3$ then L contains $\ell^{2(n+k(L))}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.*

Proof. Assume first $j_n \leq 1$. It is clear that every element of $G(\ell^n)$ can we written as $\lambda\,\mathrm{Id} + m_{\ell^n}$, where $\lambda \in \mathbb{Z}/\ell^n\mathbb{Z}$ and $m_{\ell^n} \in L(\ell^n)$. Now L is generated by x_1, x_2, x_3, so in turn every m_{ℓ^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 $l_{\ell^n} \in L(\ell^n)$ such that, for every m_{ℓ^n}, there exists a scalar $\mu \in \mathbb{Z}/\ell^n\mathbb{Z}$ with $m_{\ell^n} = \mu l_{\ell^n}$. It follows that every element of $G(\ell^n)$ can be written as $\lambda\,\mathrm{Id} + \mu l_{\ell^n}$ for suitable λ, μ, and since Id and l_{ℓ^n} commute with themselves and each other our claim follows.

Next consider the case $j_n = 2$. We can safely assume that $L(\ell^{2n})$ still has 2 generators, for otherwise we are done. It is also clear that the two generators of $L(\ell^n)$ and $L(\ell^{2n})$ are the same. Suppose first $\pi_n(x_1) = 0$, so that $k(L) \geq 1$. Then $G(\ell^n)$ is a subset of

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

and $\mathrm{Id}, \pi_n(x_2), \pi_n(x_3)$ 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

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

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

(6.1)
$$\begin{cases} -a_{21}b_{12} \equiv \xi_2 b_{11} \pmod{\ell^{2n-k(L)}} \\ 4(a_{11}b_{12} - a_{12}b_{11}) \equiv \xi_2 b_{12} \pmod{\ell^{2n-k(L)}}. \end{cases}$$

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 (6.1) we find

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

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 $L\left(2^n\right)$.

Let $\alpha = \min\{v_2(2b_{11}), v_2(b_{21})\} \leq n$ and reduce the last congruence modulo $2^{\alpha+1}$. Then $2x_1 \cdot y \equiv x_1 \cdot (2y) \equiv 0 \pmod{2^{\alpha+1}}$, 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 $y \equiv 0 \pmod{2^{\alpha+1}}$, against 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 $x_1 \cdot \tilde{y} \equiv \left(\frac{\xi_2}{2} - a_{11}\right) \tilde{y} \pmod{\ell^{n-k(L)+1-2v}}$, and now $\tilde{y} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}$ 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 $M = \begin{pmatrix} \tilde{y}_1 & 0 \\ \tilde{y}_2 & 1 \end{pmatrix}$: this is then a matrix in $\mathrm{GL}_2(\mathbb{Z}_\ell)$, 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 M 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 M is given by

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

$$= (\xi_2/2 - a_{11}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \ell^{n-k(L)+1-2v} M^{-1} w$$

$$\equiv (\xi_2/2 - a_{11}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \pmod{\ell^{n-k(L)+1-2v}}$$

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

Finally, suppose $j_n = 3$. Given our definition of a reduced basis, this means $\pi_n(x_3) \neq 0$, so $v_\ell(c_{12}) \leq n - 1$. Moreover, L contains

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

whose diagonal entries have valuation at most $v_\ell(a_{21}) + v_\ell(c_{12}) \leq n + k(L) - 1$. Consider then the matrix $\widehat{h} = \ell^{n-1} h$. The coefficient in its top-right corner, call it $\widehat{h}_{12}$, has valuation at least $n - 1$, so $\widehat{h}_{12}/c_{12}$ is in $\mathbb{Z}_\ell$, and L contains

$$x_4 = \widehat{h} - \frac{\widehat{h}_{12}}{c_{12}} x_3 = \begin{pmatrix} -\ell^{n-1} a_{21} c_{12} & 0 \\ 0 & \ell^{n-1} a_{21} c_{12} \end{pmatrix},$$

whose only nonzero coefficients are the diagonal ones, of valuation at most

$$n - 1 + v_\ell(a_{21}) + v_\ell(c_{12}) \leq (n-1) + (n + k(L) - 1) = 2n + k(L) - 2.$$

In the same exact way we can subtract from $\ell^{2n+k(L)-2} x_1$ suitable multiples of x_3 and x_4 in order to show that L contains $x_5 = \begin{pmatrix} 0 & 0 \\ \ell^{2n+k(L)-2} a_{21} & 0 \end{pmatrix}$, where the valuation of the only non-zero coefficient is at most $2n + 2k(L) - 2$. 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 three elements $x_3 = \alpha_1 s_1, x_4 = \alpha_2 s_2, x_5 = \alpha_3 s_3$ that satisfy the inequalities $v_\ell(\alpha_1) \leq n - 1, v_\ell(\alpha_2) \leq 2n + k(L) - 2, v_\ell(\alpha_3) \leq 2n + 2k(L) - 2$: since

$$\max \{ n - 1, 2n + k(L) - 2, 2n + 2k(L) - 2 \} \leq 2n + 2k(L),$$

it is clear that the module generated by x_3, x_4, x_5 contains $\ell^{2(n+k(L))} \mathfrak{sl}_2(\mathbb{Z}_\ell)$, hence a fortiori so does L. $\qquad\square$

Corollary 6.4. *Let G be a closed subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ satisfying property ($\star\star$) of theorem 4.2 (resp. $G(4) = \{\mathrm{Id}\}$ and $\det(G) \equiv 1 \pmod 8$ if $\ell = 2$). Then for every 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

$$\mathcal{B}_\ell(24n) = \left(\mathrm{Id} + \ell^{24n} \mathfrak{gl}_2(\mathbb{Z}_\ell)\right) \cap \mathrm{SL}_2(\mathbb{Z}_\ell),$$

if ℓ is odd, and it contains $\mathcal{B}_2(72n + 2)$, 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 $\ell^{6n}\mathfrak{sl}_2(\mathbb{Z}_\ell)$.

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 6.3) we have

$$L \supseteq \ell^{2(2n)+2k(L)}\mathfrak{sl}_2(\mathbb{Z}_\ell) \supseteq \ell^{6n}\mathfrak{sl}_2(\mathbb{Z}_\ell)$$

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

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

In all cases, property ($\star\star$) (resp. theorem 5.1 (i) for $\ell = 2$) now implies that G' contains $\mathcal{B}_\ell(24n)$ (resp. $\mathcal{B}_2(72n + 2)$) as claimed. $\square$

7. Application to Galois groups

We now plan to apply the above machinery to the Galois representations afforded by 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 $\mathrm{Gal}\left(\overline{K}/K\right)$ 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 4.2 (resp. theorem 5.1) we know that either G_ℓ contains a subgroup H_ℓ satisfying $[G_\ell : H_\ell] \leq 120$ (respectively $[G_\ell : H_\ell] \leq 192$ for $\ell = 2$) and the hypotheses of corollary 6.4, 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 associated with H_ℓ. The degree $[K_\ell : K]$ is then bounded by 120, 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 6.4. 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 propositions 7.1 and 7.4 below, the proofs of which are inspired by the arguments of [MW93a] and [MW89].

Proposition 7.1. *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 $\mathrm{Gal}\left(\overline{K}/K\right)$-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 2.8 we also have an isogeny $\varphi : E^* \to E$ 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 $\ell^n \mid b_0(K, E)$. $\square$

Corollary 7.2. *Let L be the Lie algebra associated with 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 $\ell^{t+1} \mid k(L)$, 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 $\ell^{t+1} \nmid b_0(K, E)$. $\square$

Corollary 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 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 let $C = b_0(K, E)$. Pick any $\alpha \in G_\ell$ whose image modulo $\ell^{1+v_\ell(C)}$ is not a scalar matrix and suppose that $G_\ell(\ell^n)$ is abelian. Consider then 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 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 end up with an endomorphism φ 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 φ kills Γ we must have $e_{11}x + e_{12}\alpha(x) = 0, e_{21}x + e_{22}\alpha(x) = 0$ for every $x \in E[\ell^n]$. Let $\eta = \min\{v_\ell(e_{ij})\}$ and suppose by contradiction $\eta < n - v_\ell(C)$. For the sake of simplicity, let us assume that 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 $\ell^{n-\eta} \mid C$, so $n - \eta \leq v_\ell(C)$, a contradiction. It follows that $\ell^{2n}\ell^{-2v_\ell(C)} \mid \ell^{2\eta} \mid \det\begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}$. Squaring this last equation we have

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

so $\ell^{2n}\ell^{-4v_\ell(C)} \mid b$ and $\ell^{2n} \mid \ell^{4v_\ell(C)}b_0(K, E \times E) \mid C^4 b_0(K, E \times E)$.

The second assertion follows immediately from the fact that ℓ is prime. $\qquad\square$

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 3.

Theorem 7.5. *Let ℓ be an odd prime (resp. $\ell = 2$) and set*

$$D(\ell) = b_0(K_\ell, E)^5 b_0(K_\ell, E \times E).$$

Suppose that $\ell^{n-v} \nmid D(\ell)$: then H'_ℓ contains $\mathcal{B}_\ell(24n)$, for odd ℓ, and it contains $\mathcal{B}_2(72n + 2)$, 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 6.4. Note that the image of $\mathrm{Gal}\left(\overline{K_\ell}/K_\ell\right)$ in $\mathrm{Aut}\, T_\ell(E)$ is exactly H_ℓ by construction. We then apply corollary 6.4 to $G = H_\ell$, assuming that $\ell^{n-v} \nmid D(\ell)$. Since $\ell^{k(L)} \mid b_0(K_\ell, E)$ by lemma 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 $\ell^{n-k(L)+1-2v} \nmid b_0(K_\ell, E)^4 b_0(K_\ell, E \times E)$. But then proposition 7.1 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. It then follows from Corollary 6.4 that $G' = H'_\ell$ contains the principal congruence subgroup $\mathcal{B}_\ell(24n)$ (resp. $\mathcal{B}_\ell(72n+2)$ for $\ell = 2$). $\qquad\square$

Corollary 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 $\mathcal{B}_\ell(24n)$.*

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 $\ell^n \nmid D(\ell)$. $\qquad\square$

Corollary 7.7. *Notation as above. The index $[\mathrm{SL}_2(\mathbb{Z}_\ell) : (H'_\ell \cap \mathcal{B}_\ell(1))]$ is of the form $A(\ell)B(\ell)$, where for every ℓ the number $A(\ell)$ divides $|\mathrm{SL}_2(\mathbb{F}_\ell)|$, and for $\ell \neq 2$ the number $B(\ell)$ is a power of ℓ dividing $\ell^{69} \cdot D(\ell)^{72}$ (resp. $B(2)$ is a power of 2 dividing $2^{222} \cdot D(2)^{216}$).*

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

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

so we are left with proving that $B(\ell) = [\mathcal{B}_\ell(1) : (H'_\ell \cap \mathcal{B}_\ell(1))]$ divides $\ell^{69} D(\ell)^{72}$ (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 $\ell^n \mid\mid D(\ell)$: then $\ell^{n+1} \nmid D(\ell)$, and therefore the above theorem implies that H'_ℓ contains $\mathcal{B}_\ell(24(n+1)) \subseteq \mathcal{B}_\ell(1)$ (resp. $\mathcal{B}_2(72(n+1)+2)$ for $\ell = 2$): the index of $\mathcal{B}_\ell(24(n+1))$ in $\mathcal{B}_\ell(1)$ is $\ell^{3(24(n+1)-1)}$, so we get

$$[\mathcal{B}_\ell(1) : (H'_\ell \cap \mathcal{B}_\ell(1))] \mid \ell^{72n+69} \mid \ell^{69} \cdot D(\ell)^{72}$$

for $\ell \neq 2$ and

$$[\mathcal{B}_2(1) : H'_2] \mid 2^{3(72n+74)} \mid 2^{222} \cdot D(2)^{216}$$

for $\ell = 2$. $\qquad\square$

Remark 7.8. Let Ψ be a natural number and $B(\ell)$ as in the above corollary: then

$$\prod_{\ell \mid \Psi} B(\ell) \mid \prod_{\ell \mid \Psi} \ell^{69} \ell^{72 v_\ell(D(\ell))} \mid \prod_{\ell \mid \Psi} \ell^{69} \ell^{72 v_\ell(D(\infty))} = \prod_{\ell \mid \Psi} \ell^{69} \cdot \prod_{\ell \mid \Psi} \ell^{72 v_\ell(D(\infty))},$$

which in turn divides $\Psi^{69} \cdot D(\infty)^{72}$. We deduce in particular the inequality

$$\prod_{\ell \mid \Psi} B(\ell) \leq \Psi^{69} D(\infty)^{72}.$$

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 8.1. *The index*

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

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

Proof. The Weil pairing induces an identification of the determinant

$$\mathrm{Gal}\left(\overline{K}/K\right) \xrightarrow{\rho_\ell} G_\ell \xrightarrow{\det} \mathbb{Z}_\ell^\times$$

with $\mathrm{Gal}\left(\overline{K}/K\right) \xrightarrow{\chi_\ell} \mathbb{Z}_\ell^\times$, where χ_ℓ denotes the ℓ-adic cyclotomic character; by Galois theory we have

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

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

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

as claimed. $\qquad\square$

The other result we will need is a surjectivity result (on SL_2) modulo ℓ for every ℓ sufficiently large: as previously mentioned, these are essentially the ideas of [MW93a] and [Mas98], in turn inspired by those of Serre.

Lemma 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 3.13, if G does not contain $\mathrm{SL}_2(\mathbb{F}_\ell)$, then the following are the only possibilities:

(I) 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 Proposition 7.1.

(II) G is contained in a Cartan subgroup of $\mathrm{SL}_2(\mathbb{F}_\ell)$: then

$$\ell \mid b_0(K, E)^4 b_0(K, E \times E) \Rightarrow \ell \mid b_0(K, E) b_0(K, E \times E)$$

by proposition 7.4.

(III) G is contained in the normalizer of a Cartan subgroup of $\mathrm{SL}_2(\mathbb{F}_\ell)$: let C be this Cartan subgroup and N its normalizer. By Dickson's classification C has index 2 in N, so the morphism $\mathrm{Gal}\left(\overline{K}/K\right) \to G \to \dfrac{G}{G \cap C} \hookrightarrow \dfrac{N}{C}$ induces a quadratic character of $\mathrm{Gal}\left(\overline{K}/K\right)$, whose kernel is associated with a certain field K' satisfying $[K' : K] \leq |N/C| = 2$. By construction, then, the image of $\mathrm{Gal}\left(\overline{K'}/K'\right)$ in $\mathrm{Aut}\left(E[\ell]\right)$ is contained in C, so applying proposition 7.4 to $E_{K'}$ we get

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

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

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

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 $\mathbb{F}_\ell^\times$: proposition 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. $\qquad\square$

Corollary 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 3.14. $\qquad\square$

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 9.1. *Let E/K be an elliptic curve without complex multiplication with stable Faltings height $h(E)$. Let*
$$\rho_\infty : \mathrm{Gal}\left(\overline{K}/K\right) \to \mathrm{GL}_2\left(\widehat{\mathbb{Z}}\right)$$
be the adelic Galois representation associated with E, and set
$$\Psi = 2 \cdot 3 \cdot 5 \cdot b_0(K, E \times E; 2) b_0(K, E; 60), \quad D(\infty) = b_0(K, E; 120)^5 b_0(K, E \times E; 120);$$
let moreover K_2 be as in section 7 and
$$D(2) = b_0(K_2, E)^5 b_0(K_2, E \times E).$$
With this notation we have
$$\left[\mathrm{GL}_2\left(\widehat{\mathbb{Z}}\right) : \rho_\infty \mathrm{Gal}\left(\overline{K}/K\right)\right] \leq [K : \mathbb{Q}] \cdot 2^{225} D(2)^{216} \cdot (\Psi D(\infty))^{72}.$$

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 $L_{cyc} = L(\mu_\infty)$ be its maximal cyclotomic extension. From the exact sequence
$$1 \to \frac{\mathrm{SL}_2(\widehat{\mathbb{Z}})}{\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)} \to \frac{\mathrm{GL}_2(\widehat{\mathbb{Z}})}{\rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K\right)\right)} \to \frac{\widehat{\mathbb{Z}}^\times}{\det \circ \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K\right)\right)} \to 1$$
we see that $[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K\right)\right)]$ equals
$$[\widehat{\mathbb{Z}}^\times : \det \circ \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K\right)\right)] \cdot [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)],$$
where the first term is bounded by $[K : \mathbb{Q}]$ thanks to proposition 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 ρ_ℓ is not surjective, and let F be the field generated over K by $\bigcup_{\ell \in \mathcal{P}} E[\ell]$. It is clear that
$$[\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)] \leq [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/F_{cyc}\right)\right)].$$

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 9.2. *Let $B(\ell)$ be as in corollary 7.7 and $D(2)$ be as in the statement of theorem 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
$$\left[\mathrm{SL}_2(\mathbb{Z}_2) : S_2\right] < 2^{225} D(2)^{216}.$$
(3) For $\ell \notin \mathcal{P}$ the equality $S_\ell = \mathrm{SL}_2(\mathbb{Z}_\ell)$ holds.

Proof. 1. This would follow from [Ser13, 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 $p_B : S \to S_B = \prod_{\ell \in B} S_\ell$ 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 $S_\ell = \rho_\ell(\mathrm{Gal}\left(\overline{F}/F\right))$ is a pro-ℓ group for every $\ell \in \mathcal{P}$: indeed, 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 S is pro-nilpotent as well and therefore it is the product of its pro-Sylow subgroups (which are just the S_ℓ's). To treat the general case we recall some terminology from [Ser97]. 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 [Ser97, 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

$$(9.1) \qquad\qquad 1 \to N_{\ell_0} \to p_{B \cup \{\ell_0\}}(S) \to p_B(S) \to 1$$

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 $\mathrm{PSL}_2(\mathbb{F}_{\ell_0}) \not\cong A_5$ since $\ell_0 \neq 5$), and therefore it must occur in N_{ℓ_0}. Denote $\overline{N_{\ell_0}}$ 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 $\mathrm{Occ}\left(N_{\ell_0}\right)$ equals $\mathrm{Occ}\left(\overline{N_{\ell_0}}\right)$ 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 [Ser97, IV-23, Lemma 2], and by lemma 3.14 this implies $N_{\ell_0} = \mathrm{SL}_2(\mathbb{Z}_{\ell_0})$: by (9.1) 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 7), so we just need to invoke corollary 7.7 to have

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

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 7.7, which now reads

$$\left[\mathrm{SL}_2(\mathbb{Z}_2) : S_2\right] \leq 2^{222} D(2)^{216} |\mathrm{SL}_2(\mathbb{F}_2)| < 2^{225} D(2)^{216}.$$

3. As $\ell \notin \mathcal{P}$ we know that $\rho_\ell(\mathrm{Gal}\left(\overline{K}/K\right))$ contains $\mathrm{SL}_2(\mathbb{Z}_\ell)$, so the group $\mathrm{PSL}_2(\mathbb{F}_\ell)$ occurs in $\rho_\ell(\mathrm{Gal}\left(\overline{K}/K\right))$. 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 $\bigcup_{p \in \mathcal{P}} \mathrm{Occ}(\mathrm{GL}_2(\mathbb{Z}_p))$, and in particular $\mathrm{PSL}_2(\mathbb{F}_\ell)$ does not occur in $\mathrm{Gal}(F/K)$. Now $\rho_\ell(\mathrm{Gal}\left(\overline{K}/K\right))$ 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 (i), we then see that S_ℓ projects surjectively onto $\mathrm{PSL}_2(\mathbb{F}_\ell)$, and therefore $S_\ell = \mathrm{SL}_2(\mathbb{Z}_\ell)$. $\qquad\square$

The proof of theorem 9.1 is now immediate:

Proof. (of theorem 9.1) We have already seen that the index $\left[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K\right)\right)\right]$ equals $[\mathbb{Z}^\times : \det \circ \rho_\infty \mathrm{Gal}\left(\overline{K}/K\right)] \cdot [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : \rho_\infty\left(\mathrm{Gal}\left(\overline{K}/K_{cyc}\right)\right)]$. 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

$$[K : \mathbb{Q}] \cdot [\mathrm{SL}_2(\widehat{\mathbb{Z}}) : S] \leq [K : \mathbb{Q}] \cdot \prod_{\ell \in \mathcal{P}} [\mathrm{SL}_2(\mathbb{Z}_\ell) : S_\ell]$$

$$\leq [K : \mathbb{Q}] \cdot \prod_{\ell | \Psi} [\mathrm{SL}_2(\mathbb{Z}_\ell) : S_\ell]$$

$$< [K : \mathbb{Q}] \cdot 2^{225} \cdot D(2)^{216} \cdot \prod_{\ell | \Psi, \ell \neq 2} |\mathrm{SL}_2(\mathbb{F}_\ell)| \cdot \prod_{\ell | \Psi, \ell \neq 2} B(\ell),$$

where we have used the fact that $\ell \nmid \Psi \Rightarrow \ell \notin \mathcal{P}$. We now bound $|\mathrm{SL}_2(\mathbb{F}_\ell)|$ with ℓ^3 and recall that (by remark 7.8) $\prod_{\ell | \Psi, \ell \neq 2} B(\ell) \leq \Psi^{69} D(\infty)^{72}$. Therefore the adelic index is at most

$$[K : \mathbb{Q}] \cdot 2^{225} \cdot D(2)^{216} \cdot \left(\prod_{\ell | \Psi, \ell \neq 2} \ell^3 \right) \cdot \Psi^{69} \cdot D(\infty)^{72},$$

which in turn is less than $[K : \mathbb{Q}] \cdot 2^{225} \cdot D(2)^{216} \cdot \Psi^{72} \cdot D(\infty)^{72}$, whence the theorem. $\square$

Using the estimates of proposition 2.6 to bound $\Psi, D(2)$ and $D(\infty)$ we get:

Corollary 9.3. *(Theorem 1.1) Let E/K be an elliptic curve that does not admit complex multiplication. The inequality*

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

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

Remark 9.4. Le Fourn has proved a much better estimate for Ψ ([Le 12], Theorem 4.2) which could be used to improve the above bound. Unfortunately, his techniques do not seem to be easily applicable to bound $D(\infty)$. Since our estimates for Ψ and $D(\infty)$ are essentially of the same order of magnitude, 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, using the improved version of the isogeny theorem mentioned in remarks 2.3 and 2.7 it is easy to see that we can take $C_1 = \exp\left(6.5 \cdot 10^{13} \right)$ and $C_2 = 24480$.

9.1. The field generated by a torsion point.
As an easy consequence of our main result we can also prove:

Corollary 1.3. *Let E/K be an elliptic curve that does not admit complex multiplication. There exists a constant $C(E/K)$ with the following property: for every $x \in E_{\mathrm{tors}}(\overline{K})$ (of order denoted $N(x)$) the inequality*

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

holds. We can take $C(E/K) = \left(\zeta(2) \cdot \left[\mathrm{GL}_2\left(\widehat{\mathbb{Z}} \right) : \rho_\infty \mathrm{Gal}\left(\overline{K}/K \right) \right] \right)^{-1}$, which is explicit 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 $\mathrm{Gal}\left(\overline{K}/K \right)$ inside $\mathrm{Aut}\, E[N]$, which we identify to $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ 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 $\mathrm{Stab}(x) = \left\{ \sigma \in G(N) \mid \sigma(x) = x \right\}$. 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 $\displaystyle\prod_{p\text{ prime}}\left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)}$. $\square$

REFERENCES

[GR14]   Éric Gaudron and Gaël Rémond. Polarisations et isogénies. *Duke Math. J.*, 163(11):2057–2108, 2014.

[Le 12]   S. Le Fourn. Surjectivity of Galois representations associated with quadratic Q-curves. *ArXiv e-prints*, 1212.4713, December 2012.

[Mas89]   D. W. Masser. Counting points of small height on elliptic curves. *Bull. Soc. Math. France*, 117(2):247–265, 1989.

[Mas98]   D. W. Masser. Multiplicative isogeny estimates. *J. Austral. Math. Soc. Ser. A*, 64(2):178–194, 1998.

[MW89]   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.

[MW93a]   D. W. Masser and G. Wüstholz. Galois properties of division fields of elliptic curves. *Bull. London Math. Soc.*, 25(3):247–254, 1993.

[MW93b]   David Masser and Gisbert Wüstholz. Isogeny estimates for abelian varieties, and finiteness theorems. *Ann. of Math. (2)*, 137(3):459–472, 1993.

[MW93c]   David Masser and Gisbert Wüstholz. Periods and minimal abelian subvarieties. *Ann. of Math. (2)*, 137(2):407–458, 1993.

[Pin93]   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.

[Ser72]   J.-P. Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15(4):259–331, 1972.

[Ser97]   J.-P. Serre. *Abelian ℓ-adic Representations and Elliptic Curves*. A. K. Peters Ltd, 3 edition, October 1997.

[Ser13]   Jean-Pierre Serre. Un critère d'indépendance pour une famille de représentations ℓ-adiques. *Comment. Math. Helv.*, 88(3):541–554, 2013.

[Zyw11]   D. Zywina. Bounds for Serre's open image theorem. *ArXiv e-prints*, 1102.4656, February 2011.