Diophantine Analysis and Linear Groups

PhD Thesis

Candidate:

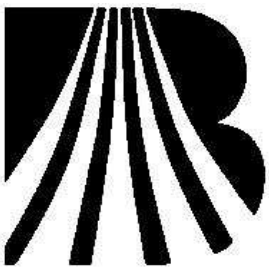
Marco
Illengo

Co-advisors:

Umberto
Zannier

Yuri
Bilu

Scuola
Normale
Superiore



Université
Bordeaux 1

Preface

The present thesis consists of two parts. The first part, which is going to appear in the Journal de Théorie des Nombres de Bordeaux, is inspired by the work of Dvornicich and Zannier [9]. They proved that for a prime p, a number field k, and a torus $\mathcal{T}$ of dimension n over k, if $n \leqslant \max\{3, 2(p-1)\}$ then the torus $\mathcal{T}$ enjoys a "local-global divisibility principle", i.e. if $P \in \mathcal{T}(k)$ is such that for almost all places ν there exists $D_\nu \in \mathcal{T}(k_\nu)$ with $pD_\nu = P$, then there exists $D \in \mathcal{T}(k)$ with $pD = P$.

Dvornicich and Zannier also showed that, when $p \neq 2$, the following condition is sufficient for the local-global divisibility:

for every p-group G in $\mathrm{SL}_n(\mathbb{Z})$ the natural map

$$\varphi\colon H^1(G, \mathbb{F}_p^n) \to \prod_C H^1(C, \mathbb{F}_p^n)$$

where C runs through the cyclic subgroups of G, is injective.

In this thesis we prove that for any prime $p \neq 2$ and any p-group G of matrices in $\mathrm{SL}_n(\mathbb{Z})$ with $n < 3(p-1)$ such injectivity automatically holds, thus extending the result of Dvornicich and Zannier. Furthermore, we prove that our result is optimal, in the sense that for $p \neq 2$ and $n \geqslant 3(p-1)$ one can always build an example where φ is not injective.

The second part is a joint work with Yuri Bilu and is inspired by his work [6]. Let X be a projective curve defined over a number field k and j a non-constant element of $k(X)$. Further, let K be a finite extension of k, let S be a finite set of places on K (which includes all the infinite places), and $\mathcal{O}_S$ the ring of S-integers of the field K. The celebrated theorem of Siegel states that if either $\mathbf{g}(X) \geqslant 1$ or j has at least 3 poles then the set of S-integral points $X(\mathcal{O}_S, j) = \{P \in X(K) \mid j(P) \in \mathcal{O}_S\}$ is finite. A couple (X, j) which satisfies these hypotesis is called "Siegelian".

The proof of Siegel's theorem does not provide any upper bound on the size of the S-integral points P of X, i.e. on the height of $j(P)$. Nonetheless, in some special cases there have been obtained "effective" versions of this theorem, which provide effective upper bounds in terms of K, S, and (X, j).

Bilu [5, 6] proved effective Siegel's theorem for some classes of modular curves, namely for (X_Γ, j) when Γ is one of the classical subgroups $\Gamma(N)$, $\Gamma_1(N)$, $\Gamma_0(N)$, provided the corresponding pair (X_Γ, j) is Siegelian.

In this thesis we prove effective Siegel's theorem for (X_Γ, j) when Γ is "almost every" congruence subgroup. In the prime power level our result is nearly best possible: our methods cover all but one case, up to equivalence. In the general case we prove effective Siegel's theorem for every Siegelian couple (X_Γ, j), provided the level of Γ does not divide a certain integer.

Prefazione

Questa tesi si divide in due parti. La prima parte, che comparirà nel Journal de Théorie des Nombres de Bordeaux, è ispirata da un lavoro di Dvornicich e Zannier [9]. Loro hanno provato che per un primo p, un campo di numeri k ed un toro $\mathcal{T}$ di dimensione n su k, se $n \leqslant \max\{3, 2(p-1)\}$ allora il toro $\mathcal{T}$ gode di un "principio di divisibilitá locale-globale", cioï£¡se $P \in \mathcal{T}(k)$ è tale che per quasi tutti i posti ν esiste $D_\nu \in \mathcal{T}(k_\nu)$ con $pD_\nu = P$, allora esiste $D \in \mathcal{T}(k)$ con $pD = P$.

Dvornicich e Zannier hanno anche mostrato che, per $p \neq 2$, la condizione seguente è sufficiente per la divisibilità locale-globale:

per ogni p-gruppo G in $\mathrm{SL}_n(\mathbb{Z})$ l'applicazione naturale

$$\varphi\colon H^1(G, \mathbb{F}_p^n) \to \prod_C H^1(C, \mathbb{F}_p^n)$$

dove C varia tra i sottogruppi ciclici di G, è iniettiva.

In questa tesi dimostriamo che per ogni primo $p \neq 2$ ed ogni p-gruppo G di matrici in $\mathrm{SL}_n(\mathbb{Z})$ con $n < 3(p-1)$ tale iniettività vale automaticamente, estendendo dunque il risultato di Dvornicich e Zannier. Inoltre, mostriamo che il nostro risultato è ottimale, nel senso che per $p \neq 2$ e $n \geqslant 3(p-1)$ si può sempre costruire un esempio per cui φ non sia iniettiva.

La seconda parte è un lavoro scritto con Yuri Bilu ed è ispirato dal suo lavoro [6]. Sia X una curva proiettiva definita su un campo di numeri k e sia j un elemento non costante di $k(X)$. Siano inoltre K un'estensione finita di k, S un insieme finito di posti su K (che includa i posti infiniti) e $\mathcal{O}_S$ l'anello degli S-interi sul campo K. Il noto teorema di Siegel afferma che se $\mathbf{g}(X) \geqslant 1$, o se j ha almeno 3 poli, allora l'insieme dei punti S-interi $X(\mathcal{O}_S, j) = \{P \in X(K) \mid j(P) \in \mathcal{O}_S\}$ è finito. Una coppia (X, j) che soddisfa queste ipotesi è detta "Siegeliana".

La dimostrazione del teorema di Siegel non fornisce alcun controllo sulla taglia dei punti S-interi P di X, cioè sull'altezza di $j(P)$. Cionondimeno, in alcuni casi particolari sono state ottenute delle versioni "effettive" del teorema, che forniscono limiti superiori in termini di K, S e (X, j).

Bilu [5, 6] ha provato un teorema di Siegel effettivo per alcune classi di curve modulari, ovvero per (X_Γ, j) dove Γ è uno dei sottogruppi classici $\Gamma(N)$, $\Gamma_1(N)$, $\Gamma_0(N)$, posto che la corrispondente coppia (X_Γ, j) sia Siegeliana.

In questa tesi dimostriamo un teorema di Siegel effettivo per (X_Γ, j) dove Γ è "quasi ogni" sottogruppo di congruenza. Nel livello potenza di primo il nostro risultato è quasi il migliore possibile: i nostri metodi trattano tutti i casi tranne uno, salvo equivalenza. Nel caso generale dimostriamo un teorema di Siegel effettivo per ogni coppia Siegeliana (X_Γ, j), posto che il livello di Γ non divida un certo intero.

Préface

Cette thèse se compose de deux parties. La première partie, qui paraitra dans le Journal de Théorie des Nombres de Bordeaux, est inspirée du travail de Dvornicich et Zannier [9]. Ils ont montré que pour un nombre premier p, un corps de nombres k et un tore $\mathcal{T}$ de dimension n sur k, si $n \leqslant \max\{3, 2(p-1)\}$ alors le tore $\mathcal{T}$ jouit du "principe de divisibilité locale-globale", c'est à dire si $P \in \mathcal{T}(k)$ est tel que pour presque toute place ν il existe $D_\nu \in \mathcal{T}(k_\nu)$ avec $pD_\nu = P$, alors il existe $D \in \mathcal{T}(k)$ avec $pD = P$.

Dvornicich et Zannier ont aussi montré que, quand $p \neq 2$, la condition suivante est suffisante pour la divisibilité locale-globale:

pour tout p-groupe G dans $\mathrm{SL}_n(\mathbb{Z})$ l'application naturelle

$$\varphi\colon H^1(G, \mathbb{F}_p^n) \to \prod_C H^1(C, \mathbb{F}_p^n)$$

où C se deplace entre le sous-groupes cycliques de G, est injective.

Dans cette thèse nous montrons que pour chaque premier $p \neq 2$ et chaque p-groupe G de matrices dans $\mathrm{SL}_n(\mathbb{Z})$ avec $n < 3(p-1)$ l'injectivité est automatique, en étendant aussi le resultat de Dvornicich et Zannier. En plus, nous montrons que notre resultat est optimal, dans le sens que pour $p \neq 2$ et $n \geqslant 3(p-1)$ on peut construire un example où φ n'est pas injective.

La deuxième partie est le resulté d'une collaboration avec Yuri Bilu et est inspirée de son travail [6]. Soit X une courbe projective définie sur un corps de nombres k et soit j un élément non constant de $k(X)$. De plus, soit K une extension finie di k, soit S un ensemble fini de places sur K (qui contient tous les places infinis) et soit $\mathcal{O}_S$ l'anneau des S-entiers sur le corps K. Le bien connu théorème de Siegel dit que si soit $\mathbf{g}(X) \geqslant 1$ soit j a au moins 3 poles alors l'ensemble des points S-entiers $X(\mathcal{O}_S, j) = \{P \in X(K) \mid j(P) \in \mathcal{O}_S\}$ est fini. Une paire (X, j) qui satisfait ces hypothèses est appelée "Siegelienne".

La démonstration du théorème de Siegel ne donne pas des bornes du haut sur la taille des points S-entiers P de X, c'est à dire sur la hauteur de $j(P)$. Néanmoins, dans des cas speciaux on a obtenu des versions "effectives" du théorème, qui donnent des bornes effectives en termes de K, S et (X, j).

Bilu [5, 6] a démontré un théorème de Siegel effectif pour certaines classes de courbes modulaires, c'est à dire pour (X_Γ, j) quand Γ est l'un des sous-groupes classiques $\Gamma(N)$, $\Gamma_1(N)$, $\Gamma_0(N)$, pourvu que la paire correspondante (X_Γ, j) soit Siegelienne.

Dans cette thèse nous démontrons un théorème de Siegel effectif pour (X_Γ, j) quand Γ est "presque quelconque" sous-groupe de congruence. Dans le niveau puissance d'un premier notre resultat est presque le meilleur possible: nos méthodes couvrent tous les cas sauf un, à équivalence près. Dans le cas général nous démontrons un théorème de Siegel effectif pour toute paire Siegelienne (X_Γ, j), pourvu que le niveau de Γ ne divise pas un certain entier.

Acknowledgements

'To absent friends, lost loves, old gods, and the season of mists; and may each and every one of us always give the devil his due.' - Hob Gadling in **Neil Gaiman's Sandman**

There are many people to whom I'm indebted for different reasons. I hope I'll be able to remember them all, and to thank them. Still, by the principle of the seventh dwarf, I am quite sure that I will forget somebody. Thus I wish to apologize in advance to the people whose name should have appeared here.

As a classical form of *captatio benevolentiæ*, I'd like to start with the members of the Jury: we could certainly say that they were an essential ingredient in the defence of the present thesis.

I can now proceed with the other acknowledgements. Being italian I refuse to, say, thank my parents in english; thus I shall mainly write in italian, with some french and english, according to the person I will be referring to.

Per primi ringrazio mamma e papà, Luciana e Mario, che hanno creato il mondo (almeno per quanto mi riguarda) e mi ci hanno instancabilmente guidato, lasciandomi libero di prendere le mie decisioni e fornendomi sempre affetto e supporto.

Un grazie a Marina, per avermi mostrato col proprio esempio che l'Italia è un'ambito molto limitato e che è persino possibile andare all'estero.

Ringrazio anche mia sorella Valentina; attraverso alti e bassi, collaborando o scontrandoci, siamo cresciuti insieme.

Ultimo, ma non per ultimo, il mio nipotino Francesco: un bimbo molto dolce, ma con una testa adatta a piantare i chiodi nel muro. Grazie per le lotte, gli agguati, gli scherzi e l'affetto.

I would like to thank my two advisors, Umberto Zannier and Yuri Bilu for the part they played in my PhD studies and thesis. From Umberto I learned that the path of a mathematician is not easy, and from Yuri I learned that one can nonetheless pretend.

Un ringraziamento generico ma sentito va ai miei "colleghi", con i quali ho avuto modo di chiacchierare diverse volte e che non ho mancato di sfruttare o tediare con vari dubbi e problemi di matematica.

Un grazie particolare a Marco Strambi per la sua disponibilità in quasi ogni ambito (*quasi*, non fraintendiamo). Una disponibilità forse eccessiva, tanto da rallentare la stesura della sua tesi (cfr [17]).

A causa del proprio argomento di tesi, una vittima perfetta a cui esporre tanti miei problemi è stata Laura Paladino: purtroppo adesso toccherà a me ricambiare.

A seguire, Anna Morra: è stato piacevole trovare degli italiani con cui fare comunella a Bordeaux, anche se questo ha significato dover mangiare *lardons* quasi ogni giorno.

L'exhaustivité des volumes de la bibliothèque spécialisée 'Alexandre Sueur' et la maîtrise du sujet de ses employés ont joué un rôle fondamental pendant mon sejour à Bordeaux.

Un'altro ringraziamento generico, questa volta rivolto a tutte le persone cui devo lo sviluppo dei miei rapporti psico-socio-affettivi e con i quali ho condiviso profonde o vacue chiacchierate, lunghe serate e bottiglie di vino. Incolperò queste ultime delle mie eventuali dimenticanze nel seguente elenco.

Scendendo nei dettagli, comincio da un amico di vecchia data quale Bruno Zori, con cui condivido il gusto per i libri ma non quello per i film e che ha iniziato una promettente carriera da *Monsù Travet*.

Continuo con altre vecchie ma inossidabili amicizie come quelle con Federica Possavino (e Bernardo) o con Christian Guerrisi. Nonostante i miei sempre più radi passaggi da Torino continuiamo a vederci e frequentarci, *Dr House* e *Nuovo Cinema Paradiso* permettendo.

Un grazie ad Harold Mancini, fanatico degli X-Men e ideatore del *barboncino fluorescente*, cui ricordo che «potrebbe andare peggio: potrebbe piovere».

Per la prima convivenza pacifica da tre anni a questa parte ringrazio Andrea Partiti, con i miei auguri per la sua ricerca del Tomo dell'Armageddon.

Devo ricordare in questa sede l'importante apporto di Lorenzo Dello Sbarba, infaticabile compagno di pause.

Suite à nos rencontres en France, je veux remercier en français Maria Marangi et Eleonora Castaldo, auxquelles je rappelle que « les roses de ma grand-mère sont aussi jaunes que mon grand-père qui était asiatique ». Merci à Maria pour les leçons sur la pensée et à Eleonora pour la théorie des *dernières trois mots*.

Un grazie anche al *principe di Firenze* Leandro Arosio, per gli insegnamenti riguardo al comportamento sportivo.

Da ultima, ma non per questo meno importante, voglio ringraziare un'oscura categoria di persone che ogni giorno lotta per salvare il mondo: quegli eroi che, chiusi nei loro angusti uffici, sventano quotidianamente i malvagi piani della burocrazia. Un grazie in particolare ad Elisabetta Terzuoli, che in quest'ultimo anno mi ha ripetutamente (chiedo scusa) *parato il culo*.

« Comment pouvez-vous lire à présent?
Il fait nuit » - Roxane
Edmond Rostand, Cyrano de Bergerac

Contents

Part I

Galois Cohomology and Local-Global Divisibility on the Torus

Chapter 1

Introduction

1.1 Algebraic tori

We briefly recall some definitions and notations concerning algebraic tori.

The multiplicative group $\mathbb{G}_m$ is an algebraic group whose k-rational points, for every field k, have the group structure of the multiplicative group of the field itself, i.e. $\mathbb{G}_m(k) \cong k^*$. For every integer n the product $\mathbb{G}_m^n$ of n copies of the multiplicative group has a natural embedding in the affine space $\mathbb{A}^{n+1}$ as the variety $(x_1 \cdot \ldots \cdot x_n \cdot y = 1)$ endowed with the group structure of componentwise multiplication. The group of automorphisms of $\mathbb{G}_m^n$ is isomorphic to the group of integral matrices $\mathrm{GL}_n(\mathbb{Z})$: in multiindex notation $x^v = x_1^{v_1} \cdot \ldots \cdot x_n^{v_n}$, to every automorphism $\phi \in \mathrm{Aut}(\mathbb{G}_m^n)$ there corresponds a matrix $M \in \mathrm{GL}_n(\mathbb{Z})$ such that $\phi(x^v) = x^{Mv}$ for every vector $v \in \mathbb{Z}^n$. With abuse of notation, we shall identify $\mathrm{Aut}(\mathbb{G}_m^n)$ with $\mathrm{GL}_n(\mathbb{Z})$. Note that this isomorphism is not canonical, but corresponds to the choice of a $\mathbb{Z}$-basis for the lattice $\mathbb{Z}^n$.

An algebraic torus of dimension n is an algebraic group defined over a number field k, isomorphic to $\mathbb{G}_m^n$ over a fixed algebraic closure $\bar{k}$ of k. The isomorphism $\varphi\colon \mathbb{G}_m^n \to \mathcal{T}$ is defined over some finite field extension K/k, which we can assume to be normal, with Galois group $\Sigma = \mathrm{Gal}(K/k)$. Every automorphism σ in the absolute Galois group $G_k = \mathrm{Gal}(\bar{k}/k)$ of k gives an isomorphism ${}^\sigma\varphi\colon \mathbb{G}_m^n \to \mathcal{T}$; by composition we obtain an element $\psi(\sigma) = \varphi^{-1} \circ {}^\sigma\varphi$ in$\mathrm{Aut}(\mathbb{G}_m^n)$. Since every automorphism of $\mathbb{G}_m^n$ is invariant under Galois action, the map $\psi\colon G_k \to \mathbb{G}_m^n(\mathbb{Z})$ is actually a homomorphism. The kernel of ψ contains $\mathrm{Gal}(\bar{k}/K)$, and its image is $\Delta = \psi(\Sigma)$. Thus every k-torus $\mathcal{T}$ defines a homomorphism ψ from G_k onto a finite subgroup Δ of $\mathrm{Aut}(\mathbb{G}_m^n) \cong \mathrm{GL}_n(\mathbb{Z})$.

Note that two k-algebraic tori $\mathcal{T}$ and $\mathcal{T}'$ whose isomorphisms $\varphi\colon \mathbb{G}_m^n \to \mathcal{T}$ and $\varphi'\colon \mathbb{G}_m^n \to \mathcal{T}'$ define the same homomorphisms $G_k \to \mathrm{GL}_n(\mathbb{Z})$ are isomorphic over k. Indeed, consider the map $\xi = \varphi' \circ \varphi^{-1}$; for every $\sigma \in G_k$ the equation ${\varphi'}^{-1} \circ {}^\sigma\varphi' = \varphi^{-1} \circ {}^\sigma\varphi$ implies ${}^\sigma\xi = \xi$. It follows that the isomorphism $\xi\colon \mathcal{T} \to \mathcal{T}'$ is defined over the base field k.

Note also that the choice of a different model for $\mathbb{G}_m^n$ gives rise to a group Δ' which is $\mathrm{SL}_n(\mathbb{Z})$-conjugate to Δ: for any automorphism ω of $\mathbb{G}_m^n$, the isomorphism $\varphi' = \varphi \circ \omega$ gives $\psi'(\sigma) = \omega^{-1}\psi(\sigma)\omega$ for every $\sigma \in \Sigma$.

Conversely, given a normal extension K/k with Galois group $\Sigma = \mathrm{Gal}(K/k)$ and a homomorphism $\psi\colon \Sigma \to \Delta < \mathrm{GL}_n(\mathbb{Z})$, there exist a k-torus $\mathcal{T}$ of dimen-

sion n and an isomorphism $\varphi\colon \mathbb{G}_m^n \to \mathcal{T}$ defined over K such that $\varphi^{-1} \circ {}^\sigma\varphi = \psi(\sigma)$ for every $\sigma \in \Sigma$.

For instance, such a torus can be constructed by defining for every $\theta \in K$ and every $v \in \mathbb{Z}^n$ the polynomials

$$S_{\theta,v}(x) = \sum_{\sigma \in \mathrm{Gal}(K/k)} \sigma^{-1}(\theta) x^{\psi(\sigma)v},$$

which satisfy the equation $S_{\theta,\psi(\sigma)v} = S_{\sigma(\theta),v}$ for any $\sigma \in \Sigma$. Let $\theta_1, \ldots, \theta_r$ be a basis of K over k and let $e_1, \ldots, e_n$ be the canonical basis of $\mathbb{Z}^n$. Then the polynomyals S_{θ_i,e_j} define a map φ of $\mathbb{G}_m^n$ in $\mathbb{A}^{rn}$, and we may define $\mathcal{T}$ as the variety $\varphi(\mathbb{G}_m^n)$ endowed with the composition law inherited from $\mathbb{G}_m^n$; by construction we have $\varphi \circ \psi(\sigma) = {}^\sigma\varphi$ for every $\sigma \in \Sigma$, and it can be easily verified that φ is injective and that $\mathcal{T}$ is actually defined over k.

Consider for instance the affine variety $(4r^3 + 2s^3 + t^3 - 6rst = 4)$ endowed with the composition map

$$\left(r_1r_2 + \frac{s_1t_2}{2} + \frac{t_1s_2}{2},\ r_1s_2 + s_1r_2 + \frac{t_1t_2}{2},\ r_1t_2 + s_1s_2 + t_1r_2 \right).$$

It is an algebraic torus $\mathcal{T}$ of dimension 2 defined over $\mathbb{Q}$. Over the normal extension $K = \mathbb{Q}(\alpha, \omega)$ of $\mathbb{Q}$, where $\alpha = \sqrt[3]{2}$ and $\omega = e^{2\pi i/3}$, we have an isomorphism $\varphi\colon \mathbb{G}_m^2 \to \mathcal{T}$ defined by

$$\varphi(x,y) = (\frac{1}{3}(x+y+x^{-1}y^{-1}),\ \frac{\alpha}{3}(x+\omega y+\omega^2 x^{-1}y^{-1}),\ \frac{\alpha^2}{3}(x+\omega^2 y+\omega x^{-1}y^{-1})).$$

Note that every $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ acts on φ as a permutation of the variables $x, y, x^{-1}y^{-1}$; in particular, the representation of $\mathrm{Gal}(K/\mathbb{Q}) \cong \S_3$ as a group of permutation matrices in $\mathrm{GL}_3(\mathbb{Z})$ contains the representation ψ of order 2.

1.2 The local-global problem

Let $\mathcal{A}$ be a commutative and connected algebraic group, with additive notation for the composition, defined over a number field k, and let m be a positive integer. We consider the following *local-global* divisibility problem.

Problem $(m, k, \mathcal{A})$ *Let $P \in \mathcal{A}(k)$ be such that for* almost all[1] *completions k_ν of k there exists a point $D_\nu \in \mathcal{A}(k_\nu)$ such that $mD_\nu = P$. Does there exist a point $D \in \mathcal{A}(k)$ such that $mD = P$?*

When $\mathcal{A}$ is the affine group $\mathbb{A}^n$, there is a well-defined k-rational map $P \mapsto \frac{1}{m}P$ and the local-global divisibility problem has a trivially positive answer. A strong form of the classical Hasse principle asserts that the local global divisibility by $m = 2$ holds on the multiplicative group $\mathcal{A} = \mathbb{G}_m$ for any number field k. Another example where the local-global divisibility problem has a positive answer is the case when $\mathcal{A} = \mathcal{E}$ is an elliptic curve and $m = p$ is a prime.

On the other hand, the problem can have a negative answer already when $\mathcal{A}$ is the multiplicative group $\mathbb{G}_m$, for instance with $k = \mathbb{Q}$, $m = 8$, and $P = 16$.

[1] By "almost all" we mean "all except possibly a finite number".

Although the problem could be put with the stronger condition that the hypothesis hold *for all* completions k_ν of k, this set-up allows us to drop the conditions for any further finite set of places ν. Obviously, if the thesis of the problem hold then its hypothesis hold for all places ν.

Dvornicich and Zannier showed in [9] that the local-global divisibility problem has a positive answer if a certain map of Galois cohomology is injective. In their subsequent paper [10], they proved that this sufficient condition is also necessary if one is allowed to enlarge the base field. We shall come back in the following chapters to these results, explaining them with more detail.

Note that the problem on the triple $(m, k, \mathcal{A})$ can be studied by investigating all cases when m is a prime power; the first natural step in this direction is the case when $m = p$ is a prime.

In [9] Dvornicich and Zannier also studied the special case when $m = p$ is a prime and $\mathcal{A} = \mathcal{T}$ is an algebraic torus of dimension n defined over a number field k; they gave a positive answer to the local-global divisibility problem on $(p, k, \mathcal{T})$ under a condition expressed in terms of p and n only, regardless of the base field k or of the precise structure of $\mathcal{T}$.

Theorem 1.2.1 *Let p be a prime and let $\mathcal{T}$ be an algebraic torus of dimension $n \leqslant \max\{3, 2(p-1)\}$ defined over a number field k. Let $P \in \mathcal{T}(k)$; if for almost all completions k_ν of k there exists a $D_\nu \in \mathcal{T}(k_\nu)$ with $pD_\nu = P$, then there exists a $D \in \mathcal{T}(k)$ such that $pD = P$.* □

For $p \neq 2$ and with n as above, they reduced the proof of this theorem to a result on group cohomology.

Proposition 1.2.2 *Let $p \neq 2$ be a prime, let $n \leqslant 2(p-1)$ be a positive integer, and let G be a p-group in $\mathrm{SL}_n(\mathbb{Z})$. Then the natural map of cohomology groups*

$$H^1(G, A) \to H^1(C, A),$$

where C runs among all cyclic subgroups of G and $A = \mathbb{Z}/(p)^n$, is injective. □

Dvornicich and Zannier suggested that this proposition could hold with less restraints on n, so to give more precise answers to the following problem of local-global divisibility on algebraic tori.

Problem (p, n) *Let p be a prime and n be a positive integer. Under which conditions on p and n does the local-global divisibility by p hold on every algebraic torus $\mathcal{T}$ of dimension n defined over a number field k?*

Note that the problem does not involve any condition on the number field k. It simply asks whenever, for fixed p and n, there exists or not an algebraic point on some algebraic torus that provides a counterexample for the local-global divisibility by p.

As Dvornicich and Zannier pointed out in [10], this problem is not trivial, i.e. some condition on the dimension of the torus is actually necessary.

Theorem 1.2.3 *Let p be a prime and let $n \geqslant p^4 - p^2 + 1$ be an integer. There exist an algebraic torus $\mathcal{T}$ of dimension n defined over a number field k and a point $P \in \mathcal{T}(k)$ such that for almost all completions k_ν of k there exists a $D_\nu \in \mathcal{T}(k_\nu)$ with $pD_\nu = P$, yet there exist no $D \in \mathcal{T}(k)$ with $pD = P$.* □

The results of Theorems 1.2.1 and 1.2.3 provide a partial answer to the above problem: the local-global divisibility by a prime p holds on every torus whose dimension is 'small enough', namely $n \leqslant \max\{3, 2(p-1)\}$, but fails for at least one torus of 'higher' dimension, i.e. $n = p^4 - p^2 + 1$. Nonetheless, for a torus of 'intermediate' dimension, these theorems do not provide any information for the local-global divisibility by p, leaving a gap of uncertainty on n.

1.3 Main results

In this part of the thesis we completely answer the question on (p, n) of local-global divisibility on algebraic tori when $p \neq 2$, namely we determine the precise bound: for any odd prime p, every algebraic torus which does not enjoy the local-global divisibility by p over some number field has dimension $n \geqslant 3(p-1)$, and the equality is attained in at least one case.

Theorem 1.3.1 *Let $p \neq 2$ be a prime and let $\mathcal{T}$ be an algebraic torus of dimension $n < 3(p-1)$ defined over a number field k. Let $P \in \mathcal{T}(k)$; if for almost all completions k_ν of k there exists a $D_\nu \in \mathcal{T}(k_\nu)$ with $pD_\nu = P$, then there exists a $D \in \mathcal{T}(k)$ such that $pD = P$.*

Theorem 1.3.2 *Let $p \neq 2$ be a prime and let $n \geqslant 3(p-1)$ be an integer. There exist an algebraic torus $\mathcal{T}$ of dimension n defined over a number field k and a point $P \in \mathcal{T}(k)$ such that the equation $pD = P$ has a solution $D_\nu \in \mathcal{T}(k_\nu)$ for almost all completions k_ν of k, but no solution $D \in \mathcal{T}(k)$.*

More precisely, exploiting some results of Dvornicich and Zannier from [9] and [10], we find more precise conditions under which the thesis of Proposition 1.2.2 holds. We determine a weaker condition on the positive integer n in terms of the odd prime p.

Theorem 1.3.3 *Let $p \neq 2$ be a prime and let $n < 3(p-1)$ be a positive integer. For every finite p-group G in $\mathrm{SL}_n(\mathbb{Z})$ the map $H^1(G, A) \to \prod H^1(C, A)$, where $A = \mathbb{F}_p^n$ and the product is taken over all cyclic subgroups C of G, is injective.*

We also show that our condition on n is 'best possible'.

Theorem 1.3.4 *Let $p \neq 2$ be a prime and $n \geqslant 3(p-1)$ be an integer. There exists a finite p-group G in $\mathrm{SL}_n(\mathbb{Z})$ such that the map $H^1(G, A) \to \prod H^1(C, A)$, where $A = \mathbb{F}_p^n$ and the product is taken over all cyclic subgroups C of G, is not injective.*

In Chapter 2 we repeat some of the arguments of [9], showing that Theorem 1.3.1 follows from Theorem 1.3.3, which we subsequently prove using some results from the geometry of numbers and from the theory of linear representations.

In Chapter 3 we briefly resume some results from [9] and [10], showing that Theorem 1.3.2 is inferred by Theorem 1.3.4; we prove the latter by constructing a counterexample in the case $n = 3(p-1)$, obtaining the general case by means of a direct sum with the trivial representation of dimension $n - 3(p-1)$.

Throughout this thesis we shall denote by I the identity matrix and by O the null matrix, whenever their orders are known.

Chapter 2

Local-global divisibility

In this chapter we prove Theorem 1.3.1. We first show how to reduce it to Theorem 1.3.3, then we prove the latter. Although the first step is described in full detail in [9], for the sake of completeness we shall briefly resume the arguments and prove the results that are applied in both this and the following chapter.

2.1 Group cohomology and divisibility

Since the arguments we shall describe involve some (first level) group cohomology, we begin by recalling some definitions for the reader's convenience.

Let A be an abelian group and let G be a group acting on A; for any $g \in G$ and any $a \in A$ we denote the image of a under the action of g by $g \cdot a = g(a)$. A *cocycle* (for the pair (G, A)) is a map $f\colon G \to A$ that satisfies the *cocycle relation* $f(gh) = f(g) + g \cdot f(h)$ for any $g, h \in G$, and a *coborder* is a map $f\colon G \to A$ such that there exists some $a \in A$ with $f(g) = g \cdot a - a$ for any $g \in G$. Every coborder satisfies the cocycle relation and actually the (additive) group $B^1(G, A)$ of all coborders is a subgroup of the group $Z^1(G, A)$ of all cocycles, and their quotient $H^1(G, A) = Z^1(G, A)/Z^1(G, A)$ is the (first) cohomology group of the pair (G, A).

Keeping our notations consistent with [9], we say that a cocycle f *satisfies the local conditions* if for every $g \in G$ there exists an $a_g \in A$ such that $f(g) = g \cdot a_g - a_g$; we also denote by $H^1_{\mathrm{loc}}(G, A)$ the image of all such cocycles in $H^1(G, A)$. Note that $H^1_{\mathrm{loc}}(G, A)$ is the kernel of the natural map

$$H^1(G, A) \to \prod H^1(C, A),$$

defined as the product of the restriction maps $H^1(G, A) \to H^1(C, A)$, where C runs among all cyclic subgroups C of G; in other words, $H^1_{\mathrm{loc}}(G, A)$ is the intersection of the kernels of all such maps.

Now, let us come back to the local-global divisibility problem in its generality: let m be an integer, let k be a number field, and let $\mathcal{A}$ be a commutative and connected algebraic group defined over k, with additive notation for the composition. We consider the (finite) set $A = \mathcal{A}[m]$ of all m-torsion points in $\mathcal{A}$ and the field $K = k(A)$, generated over k by all points in A. Note that K/k is a normal extension; we denote by $\Sigma = \mathrm{Gal}(K/k)$ its Galois group.

Proposition 2.1.1 *With the above notation, to every point $P \in \mathcal{A}(k)$ there corresponds an element $c_P \in H^1(\Sigma, A)$ with the properties:*

i) P satisfies the assumptions of the local-global divisibility problem $(m, k, \mathcal{A})$ if and only if $c_P \in H^1_{\mathrm{loc}}(\Sigma, A)$;

ii) P satisfies the conclusion of the local-global divisibility problem $(m, k, \mathcal{A})$ if and only if $c_P = 0$.

This proposition implies that the condition $H^1_{\mathrm{loc}}(\Sigma, A) = 0$ ensures a positive solution to the local global divisibility problem $(m, k, \mathcal{A})$. More precisely it can be shown that, when $m = p^e$ is a prime power, a sufficient condition is $H^1_{\mathrm{loc}}(\Sigma_p, A) = 0$, where Σ_p is the p-Sylow subgroup of Σ.

On the contrary, the condition $H^1_{\mathrm{loc}}(\Sigma, A) \neq 0$ does not necessarily ensure the existence of a k-rational point P with $c_P \neq 0$, i.e. the existence of a counterexample to the local-global divisibility by m on $\mathcal{A}$ over k. In Chapter 3 we shall see how the condition $H^1_{\mathrm{loc}}(\Sigma, A) \neq 0$ can be exploited in order to obtain some counterexample to the local-global divisibility.

We present a proof of the above proposition, referring the reader to [9] for further details.

Proof of Proposition 2.1.1 – Let $D \in \mathcal{A}(\bar{k})$ be any solution to $mD = P$ and let $L = K(D)$ be the field generated by D over K; it is a finite normal extension of k, with Galois group $\Sigma_L = \mathrm{Gal}(L/k)$.

Note that any other solution $D' \in \mathcal{A}(\bar{k})$ to $mD' = P$ differs from D by an m-torsion point, i.e. $D' - D \in A$. In particular, all solutions D' lie in $\mathcal{A}(L)$. Moreover, since P is a k-rational point, for every $\sigma \in \Sigma_L$ we $m\sigma(D) = P$, which implies that $\sigma(D) - D \in A$.

We define a map Z from Σ_L to $\mathcal{A}$ as

$$Z_\sigma = \sigma(D) - D;$$

it is immediately verified that Z is a (Σ_L, A)-cocycle: for every $\sigma, \tau \in \Sigma_L$ we have

$$Z_{\sigma\tau} = \sigma\tau(D) - \sigma(D) + \sigma(D) - D = \sigma(Z_\tau) + Z_\sigma.$$

Note that by choosing in place of D any other solution $D' = D + E$ to $mD' = P$, where $E \in A$, we would have defined a different map $Z'_\sigma = \sigma(D') - D'$. Nonetheless, in this case we would have

$$Z'_\sigma - Z_\sigma = \sigma(D' - D) - (D' - D) = \sigma(E) - E,$$

where $\sigma(E) - E$ is a (Σ_L, A)-coborder. This implies that the residue class $c_P = [Z]$ in $H^1(\sigma, A)$ depends only on P, and that the map $P \mapsto c_P$ is well-defined. Actually, every cocycle in the class $[Z]$ is obtained from some solution D' to $mD' = P$.

Since all points in A are K-rational, the action on A of the absolute Galois group $G_k = \mathrm{Gal}(\bar{k}/k)$ of k factors through Σ. This implies that for any (G_k, A)-cocycle f and any two automorphisms $\sigma, \tau \in G_k$ which extend the same element of Σ, we have $f(\sigma) = f(\tau)$; in particular, we can identify the groups $H^1(\Sigma_L, A)$ and $H^1(\Sigma, A)$.

We remark that, by definition of c_P, we have $c_P = 0$ if and only if there exists some $D' \in \mathcal{A}(L)$ which satisfies $mD' = P$ and which is invariant under

the Galois action of Σ_L, i.e. a k-rational solution D' in $\mathcal{A}$ to $mD' = P$. This proves the second part of the proposition.

Let now ν be any place on k which does not ramify in L; we can embed L in a finite extension L_w of k_ν, where w is a place of L lying above ν. The group $C_w = \mathrm{Gal}(L_w/k_\nu)$ is cyclic, generated by a Frobenius automorphism of ν relative to the field extension L/k, and it is a subgroup of Σ_L. By the same arguments as above, the existence of a k_ν-rational point D_ν on $\mathcal{A}$ such that $mD_\nu = P$ is equivalent to the vanishing of the restriction of c_P to $H^1(C_\nu, A)$.

By Čebotarev theorem, as ν runs among almost all unramified places ν of k, the group $\mathrm{Gal}(L_w/k_\nu)$ varies among all cyclic subgroups of Σ_L. This implies that P satisfies the hypotesis of the local-global divisibility problem if and only if for every cyclic subgroup C of Σ_L the restriction of c_P to $H^1(C, A)$ vanishes, i.e. if and only if $c_P \in H^1_{\mathrm{loc}}(\Sigma_L, A)$. As we have seen, this happens precisely when c_P belongs to $H^1_{\mathrm{loc}}(G, A)$. □

When restricting ourselves to the special case when $m = p \neq 2$ is a prime and $\mathcal{A} = \mathcal{T}$ is an algebraic torus of dimension n, we can exploit the isomorphism $\varphi\colon \mathbb{G}_m^n \to \mathcal{T}$, which induces a group isomorphism $\mathcal{T}[p] \cong \mathbb{G}_m^n[p] \cong \mathbb{F}_p^n$. More precisely, we shall read the action of the absolute Galois group $G_k = \mathrm{Gal}(\bar{k}/k)$ of k on $\mathcal{T}[p]$ as an action on $\mathbb{F}_p^n$. We recall that the map $\psi(\sigma) = \varphi^{-1} \circ {}^\sigma\varphi$ defines a homomorphism from G_k to a subgroup G of $\mathrm{SL}_n(\mathbb{Z}) \cong \mathrm{Aut}(\mathbb{G}_m^n)$. Letting $\chi\colon G_k \to \mathbb{F}_p^*$ be the character defined by ${}^\sigma\zeta = \zeta^{\chi(\sigma)}$, where ζ is any primitive p-th root of 1, the action of any $\sigma \in G_k$ on $\mathcal{T}[p]$ corresponds to the action on $\mathbb{F}_p^n$

$$v \mapsto \chi(\sigma)\psi(\sigma)v.$$

Let now $\xi\colon G_k \to \mathrm{GL}_n(\mathbb{F}_p)$ be the homomorphism defined by

$$\sigma \mapsto \chi(\sigma)\psi(\sigma) \pmod{p};$$

the kernel of ξ has fixed field $K = k(\mathcal{T}[p])$. We denote by $\Delta = \xi(G_k)$ its image.

By Proposition 2.1.1 and the following remarks, the local-global divisibility by p on $\mathcal{T}$ holds if $H^1_{\mathrm{loc}}(\Delta_p, \mathbb{F}_p^n) = 0$, where Δ_p is the p-Sylow subgroup of Δ, i.e. it is a p-group in $\mathrm{GL}_n(\mathbb{F}_p)$. We can actually obtain a stronger result.

Let us consider the normal extension $k(\zeta)$ of k, where ζ is a primitive p-th root of 1, and its absolute Galois group $G_{k(\zeta)}$. Restricting to $G_{k(\zeta)}$ we obtain a normal subgroup $G' = \psi(G_{k(\zeta)})$ of G and a normal subgroup $\Delta' = \xi(G_{k(\zeta)})$ of Δ. Moreover, the restriction of χ to $G_{k(\zeta)}$ is identically 1, so that G' and Δ' are isomorphic,[1] and so are their p-Sylow subgroups. Since $[k(\zeta) : k]$ is coprime with p, then so are $[G : G']$ and $[\Delta : \Delta']$; in particular, the p-Sylow subgroups G_p of G and Δ_p of Δ are isomorphic. This implies that the local-global divisibility by p on $\mathcal{T}$ is ensured by the vanishing of $H^1_{\mathrm{loc}}(G_p, \mathbb{F}_p^n)$, where G_p is some p-group in $\mathrm{SL}_n(\mathbb{Z})$.

This result allows us to obtain some information for the local-global divisibility problem on $(p, k, \mathcal{T})$ by considering any possible candidate for G_p.

Proposition 2.1.2 *Let $p \neq 2$ be a prime and let n be a positive integer. If $H^1_{\mathrm{loc}}(G_p, \mathbb{F}_p^n) = 0$ for every p-group G_p in $\mathrm{SL}_n(\mathbb{Z})$, then the local-global divisibility by p holds on every algebraic torus $\mathcal{T}$ of dimension n.*

[1] Under the condition $p \neq 2$: for $p = 2$ the projection $\Delta' \to G'$ modulo p could have a non-trivial kernel.

Conversely, the existence of a p-group G_p in $\mathrm{SL}_n(\mathbb{Z})$ with $H^1_{\mathrm{loc}}(G_p, \mathbb{F}_p^n) \neq 0$ allows us to construct a torus $\mathcal{T}$, as we have seen in Chapter 1, such that $H^1_{\mathrm{loc}}(\Sigma, A) \neq 0$. As we have said, this does not necessarily imply the existence of a k-rational point P in $\mathcal{T}$ which contradicts the local-global divisibility by p. Nonetheless, in Chapter 3 we shall use this torus to obtain a counterexample.

Dvornicich and Zannier proved the above proposition in [9] in order to infer Theorem 1.2.1 from Proposition 1.2.2. We use it to reduce Theorem 1.3.1 to Theorem 1.3.3.

2.2 Proof of Theorem 1.3.3

In this section we prove Theorem 1.3.3. We can drop every notation on fields and tori, for we shall exclusively work with p-groups of matrices acting on vector spaces.

Since the groups of matrices involved by Proposition 2.1.2 are defined over $\mathbb{Z}$, and not only modulo p, we have more restraints on their structure. We begin with the following result, which is slightly more general than needed.

Lemma 2.2.1 *Let p be a prime and let G be a p-group of matrices in $\mathrm{SL}_n(\mathbb{Q})$. If $n < p(p-1)$ then G is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^b$, for some $b \leqslant n/(p-1)$.*

Proof – Note that any non-trivial element g of G is a matrix of multiplicative order p^m, for some positive integer m. This implies that at least one of the eigenvalues of g is a p^m-th primitive root of unity; recalling that g is defined over $\mathbb{Q}$, we obtain that also any other p^m-th primitive root of unity must be an eigenvalue of g. Then the number of eigenvalues of g, bounded by its order $n < p(p-1)$, is at least $\psi(p^m) = p^{m-1}(p-1)$. It follows that $m = 1$, i.e. that g has order p. Thus G has exponent p.

Let now K be $(\mathbb{Z}/p\mathbb{Z})^*$; we say that two elements, g and h, of G are *K-conjugate* if there exists a $k \in K$ such that g^k and h are conjugate by an element of G. By the theory of characters for finite representations (see [14, Section 12.3]), the number of representations of G which are irreducible over $\mathbb{Q}$ is equal to the number of K-conjugacy classes of G. It is also well-known that the number of $\mathbb{C}$-irreducible representations of G is equal to the number of conjugacy classes of G, and that these representations can all be realized over the p-th cyclotomic field $\mathbb{Q}(\zeta_p)$.

We first enumerate the number of K-conjugacy classes of G, with respect to the number of its conjugacy classes. Assume that a non-trivial element g of G is conjugate to g^k, for some $k \in K$. This means that there exists an element h in G such that conjugation by h maps g to g^k. In this case conjugation by h^p maps g to $g^{k^p} = g^k$, since g has order p; on the other hand h^p is the neuter element, thus $g^k = g$. This shows that any two distinct powers of a same element are not conjugate, and that every K-conjugacy class of G (apart from the class of the identity element) is the union of $p-1$ distinct conjugacy classes of G. In other words, every $\mathbb{Q}$-irreducible representation of G is equivalent to the direct sum of the distinct conjugates of some $\mathbb{C}$-irreducible representation of G.

Now, if the group G was non-commutative, its faithful representation G would contain an irreducible representation of degree $d > 1$, with d dividing the order of G, i.e. with $d \geqslant p$. If this was the case, then G would also contain

a $\mathbb{Q}$-irreducible representation of degree $(p-1)d \geqslant (p-1)p > n$, which is not possible. This implies that G is an abelian group.

By the classification of abelian groups, we obtain that G is isomorphic to the direct product of b copies of $\mathbb{Z}/p\mathbb{Z}$, for some integer b. Note that any faithful representation of G over $\mathbb{C}$ has order at least b, and that any faithful representation of G over $\mathbb{Q}$ has order at least $b(p-1)$. Then $b \leqslant n/(p-1)$. □

For the rest of this chapter, we shall assume the hypotesis of Theorem 1.3.3. Thus $p \neq 2$ will be a prime number, $n < 3(p-1)$ will be an integer, and G will be a p-group of integer matrices in $\mathrm{SL}_n(\mathbb{Z})$.

We remark that, when G is a cyclic group, the theorem is trivially true. Applying Lemma 2.2.1, we obtain that the group G is cyclic (and the theorem is proved), unless $2(p-1) \leqslant n < 3(p-1)$ and $G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. Let us put ourselves in this case.

Note that the proof of Lemma 2.2.1 shows that the representation G is the direct sum of two distinct $\mathbb{Q}$-irreducible representations of order $p-1$ and $(n-2(p-1))$ copies of the trivial representation.

We remark that, extending constants to the p-th cyclotomic field $\mathbb{Q}(\zeta_p)$ and after a base-change, the representation G could be written in diagonal form, as a direct sum of its irreducible subrepresentations. Also, extending constants to $\mathbb{Q}$ and after a base-change, the representation G could be written as a direct sum of its $\mathbb{Q}$-irreducible subrepresentations. Since we are dealing with the action of G on $\mathbb{F}_p^n$, though, we shall restrict to base-changes to $\mathbb{Z}$, which are preserved under reduction modulo p.

Consider the lattice $\mathsf{N} = \mathbb{Z}^n$. It contains a sublattice $\mathsf{M} = \mathsf{N}^G$ which is fixed by G: it is the intersection of N with the subspace $(\mathbb{Q}^n)^G$ of vectors which are invariant by G. We fix a $\mathbb{Z}$-basis for M and we apply a result on lattices (see [7, Cor. 3 to Thm. 1, Ch. 1]) to extend it to a basis of N: in this way we split the lattice as $\mathsf{N} = \mathsf{M} \oplus \mathsf{L}$. Now, let ρ be one of the two non-trivial, $\mathbb{Q}$-irreducible subrepresentations of G, and let H be its kernel. Repeating the above argument on the restriction of H to L and on the sublattice L^H, we determine a basis for $\mathbb{Z}^n$ that allows us to write N in the form $\mathsf{N}^{(1)} \oplus \mathsf{N}^{(2)} \oplus \mathsf{N}^{(3)}$. Using this new basis, we can assume that every element g of G is of the form

$$g = \begin{pmatrix} I & A_g & B_g \\ O & M_g & C_g \\ O & O & N_g \end{pmatrix},$$

where $g \mapsto M_g$ and $g \mapsto N_g$ are the two $\mathbb{Q}$-irreducible representations of G of order $p-1$. In particular, we can choose generators σ and τ for G of the forms

$$\sigma = \begin{pmatrix} I & A_\sigma & B_\sigma \\ O & M & C_\sigma \\ O & O & I \end{pmatrix}; \qquad \tau = \begin{pmatrix} I & A_\tau & B_\tau \\ O & I & C_\tau \\ O & O & N \end{pmatrix}. \tag{2.1}$$

Note that the eigenvalues of M are the $p-1$ distinct p-th roots of unity. This implies that the minimal polynomial of M is $(x^p-1)/(x-1)$ and that the determinant of $M-I$ is p.

Over $\mathbb{F}_p$, the matrix M solves the polynomial $(x-1)^{p-1}$, and its minimal polynomial is thus of the form $(x-1)^s$, for some $s < p$. This implies that $(M-I)^s$ has all entries in $p\mathbb{Z}$, so that p divides every column of $(M-I)^s$.

Then p^{p-1} divides its determinant, $\det(M-I)^s = p^s$, which implies $s \geqslant p-1$. It follows that, over $\mathbb{F}_p$, the minimal polynomial of M is $(x-1)^{p-1}$ and that the Jordan form of M is a Jordan block. In particular we deduce the following proposition.

Proposition 2.2.2 *Let M be as above. For every two non-negative integers i and j with $i+j=p-1$, the image of $(M-I)^i$ is the kernel of $(M-I)^j$, i.e. for every vector A in $\mathbb{Z}^{p-1}$*

$$(M-I)^j A \equiv O \pmod{p} \iff \exists B \in \mathbb{Z}^{p-1} \mid A \equiv (M-I)^i B \pmod{p}$$

and

$$A^t(M-I)^j \equiv O \pmod{p} \iff \exists B \in \mathbb{Z}^{p-1} \mid A^t \equiv B^t(M-I)^i \pmod{p}.$$

The same holds using N in place of M. □

Note that this proposition trivially applies also to integer matrices $m \times (p-1)$, for every positive integer m.

We can now prove our theorem.

Proof of Theorem 1.3.3 – As in the above discussion, by Lemma 2.2.1 we can assume $2(p-1) \leqslant n < 3(p-1)$ and $G \cong \mathbb{Z}/(p) \times \mathbb{Z}/(p)$. Let also σ and τ be generators for G as in (2.1). A direct computation of $\sigma\tau = \tau\sigma$ gives

$$\sigma\tau = \begin{pmatrix} I & A_\sigma & \star \\ O & M & C_\sigma + C_\tau \\ O & O & N \end{pmatrix}$$

and the relations

$$A_\tau = O, \qquad (M-I)C_\tau = -C_\sigma(N-I), \qquad B_\sigma = A_\sigma(M-I)^{-1}C_\sigma. \quad (2.2)$$

Let now $\tilde{Z}$ be a $(G, \mathbb{F}_p^n)$-cocycle that satisfies the local conditions, i.e. for every g in G there exists a $\tilde{W}_g$ in $\mathbb{F}_p^n$ such that $\tilde{Z}_g = g\tilde{W}_g - \tilde{W}_g$. We choose representants W_g of $\tilde{W}_g$ in $\mathbb{Z}^n$ and we define $Z_g = gW_g - W_g$ for every g in G. Note that $\tilde{Z}_g \equiv Z_g \pmod{p}$ for every g in G.

Modulo a coboundary, we can assume $Z_\tau \equiv O \pmod{p}$. By the cocycle relation, this implies $Z_{\sigma\tau} \equiv Z_\sigma + \sigma Z_\tau \equiv Z_\sigma \pmod{p}$. By definition, Z_σ and $Z_{\sigma\tau}$ are of the form:

$$\begin{pmatrix} Z_\sigma^{(1)} \\ Z_\sigma^{(2)} \\ Z_\sigma^{(3)} \end{pmatrix} = \begin{pmatrix} A_\sigma W_\sigma^{(2)} + B_\sigma W_\sigma^{(3)} \\ (M-I)W_\sigma^{(2)} + C_\sigma W_\sigma^{(3)} \\ O \end{pmatrix},$$

$$\begin{pmatrix} Z_{\sigma\tau}^{(1)} \\ Z_{\sigma\tau}^{(2)} \\ Z_{\sigma\tau}^{(3)} \end{pmatrix} = \begin{pmatrix} \star \\ (M-I)W_{\sigma\tau}^{(2)} + (C_\sigma + C_\tau)W_{\sigma\tau}^{(3)} \\ (N-I)W_{\sigma\tau}^{(3)} \end{pmatrix};$$

where $W_\sigma^{(2)}$, $W_\sigma^{(3)}$, $W_{\sigma\tau}^{(2)}$, and $W_{\sigma\tau}^{(3)}$ all lie in $\mathbb{Z}^p$.

Since $Z_\sigma^{(3)} \equiv Z_{\sigma\tau}^{(3)} \pmod p$, we obtain $(N-I)W_{\sigma\tau}^{(3)} \equiv O \pmod p$; by Proposition 2.2.2 this implies $W_{\sigma\tau}^{(3)} \equiv (N-I)^{p-2}\tilde{R} \pmod p$, for some $\tilde{R}$ with entries in $\mathbb{F}_p$. It follows that, modulo p, the vector $(M-I)^{p-2}Z_{\sigma\tau}^{(2)}$ is of the form

$$(M-I)^{p-1}W_{\sigma\tau}^{(2)} + (M-I)^{p-2}(C_\sigma + C_\tau)(N-I)^{p-2}\tilde{R}.$$

We apply the second relation from (2.2) and the fact that both M and N solve $(x-1)^{p-1}$ modulo p, obtaining $(M-I)^{p-2}Z_{\sigma\tau}^{(2)} \equiv O \pmod p$. Applying Proposition 2.2.2 to $Z_\sigma^{(2)}$ we obtain $Z_\sigma^{(2)} \equiv (M-I)\tilde{S} \pmod p$, for some $\tilde{S}$ with entries in $\mathbb{F}_p$. Let S be any representant of $\tilde{S}$ over $\mathbb{Z}$; since the entries of $Z_\sigma^{(2)} - (M-I)S$ are all divisible by p and since $(M-I)$ has determinant p, we may assume that S satisfies $Z_\sigma^{(2)} = (M-I)S$. Thus

$$Z_\sigma^{(1)} = A_\sigma(M-I)^{-1}Z_\sigma^{(2)} = A_\sigma S.$$

In particular, we have

$$\begin{pmatrix} O & A_\sigma & B_\sigma \\ O & M-I & C_\sigma \\ O & O & O \end{pmatrix}\begin{pmatrix} O \\ S \\ O \end{pmatrix} = \begin{pmatrix} Z_\sigma^{(1)} \\ Z_\sigma^{(2)} \\ Z_\sigma^{(3)} \end{pmatrix}; \quad \begin{pmatrix} O & O & B_\tau \\ O & O & C_\tau \\ O & O & N-I \end{pmatrix}\begin{pmatrix} O \\ S \\ O \end{pmatrix} = \begin{pmatrix} O \\ O \\ O \end{pmatrix}.$$

This implies that $Z_\sigma = \sigma V - V$ and $Z_\tau \equiv \tau V - V \pmod p$, with $V = \begin{pmatrix} O \\ S \\ O \end{pmatrix}$ and, since up to a coboundary $\tilde{Z}$ vanishes on the generators σ and τ of G, that $\tilde{Z}$ itself is a $(G, \mathbb{F}_p^n)$-coboundary.

By the arbitrarity of the choice of $\tilde{Z}$, this proves that on the pair $(G, \mathbb{F}_p^n)$ any cocycle which satisfies the local conditions is a coborder, i.e. that $H^1_{\mathrm{loc}}(G, \mathbb{F}_p^n)$ is trivial, concluding the proof of Theorem 1.3.3. □

Chapter 3

A counterexample

In this chapter we prove Theorem 1.3.2. We show how to apply the arguments of Dvornicich and Zannier from [9] (mainly appearing in the previous chapter) and [10] in order to reduce it to Theorem 1.3.4, then we construct an explicit example for the latter, in the case $n = 3(p-1)$; we recall that the general case can be obtained by means of a direct sum with the trivial representation of dimension $n - 3(p-1)$.

Once again, for the sake of completeness we shall reproduce here some arguments from the above articles, referring the reader to them for more precise statements and more complete proofs.

We shall also refer to the previous chapter for some results and notations.

3.1 Construction

We recall that in Proposition 2.1.1 we considered an integer m and a commutative and connected algebraic group $\mathcal{A}$ defined over a number field k, expressing a sufficient condition for the local-global divisibility by m on $\mathcal{A}$ over k as $H^1_{\mathrm{loc}}(\Sigma, A) = 0$, where $A = \mathcal{A}[m]$, $K = k(A)$, and $\Sigma = \mathrm{Gal}(K/k)$.

As we have anticipated in the previous chapter, the condition $H^1(\Sigma, A) \neq 0$ does not necessarily ensure the existence of a counterexample to the local-global divisibility by p on $\mathcal{A}$ over k, i.e. of a k-rational point P on $\mathcal{A}$ such that the equation $mD = P$ has a k_ν-rational solution in $\mathcal{A}$ for almost all completions k_ν of k, but not k-rational solutions.

As Dvornicich and Zannier pointed out in [9], this could be the case under other conditions, for instance if $H^1(\Sigma, \mathcal{A}(K)) = 0$: in this case let Z be a (Σ, A)-cocycle which is not a (Σ, A)-coborder but satisfies the local conditions; then we have $Z \in Z^1(\Sigma, \mathcal{A}(K)) = B^1(\Sigma, \mathcal{A}(K))$, i.e. there exists some $D \in \mathcal{A}(K)$ such that $Z_\sigma = \sigma(D) - D$ for every $\sigma \in \Sigma$; as Z has values in A, for every $\sigma \in \Sigma$ we have $\sigma(mD) = mD$, which implies that $P = mD$ is a k-rational point and that it provides us the counterexample we seek.

Another case when we can exploit the non-vanishing of $H^1(\Sigma, A)$ is if we are allowed to enlarge the base field k. More precisely, Dvornicich and Zannier proved in [10] that under this condition we can find a finite extension L of k such that the local-global divisibility problem on $(m, L, \mathcal{A})$ has a negative answer.

Theorem 3.1.1 *With the above notation, if $H^1(\Sigma, A) \neq 0$ then there exist a number field L such that $L \cap K = k$ and a point $P \in \mathcal{A}(L)$ such that for almost all places w of L there exists an L_w-rational solution D_w on $\mathcal{A}$ to $mD_w = P$, but there exist no L-rational solution D on $\mathcal{A}$ to $mD = P$.*

It goes beyond the scopes of this thesis to give a complete proof of this theorem. We shall nevertheless show how it can be deduced from the following result, which Dvornicich and Zannier proved by applying Hilbert's irreducibility theorem to a previous result of Lang and Tate.

Proposition 3.1.2 *Let Z be a (Σ, A)-cocycle which satisfies the local conditions. There exists infinitely many finite extensions L of k with $L \cap K = k$ such that Z vanishes in $H^1(\Sigma, \mathcal{A}(LK))$. Moreover, one can choose infinitely many L such that the extension LK/L is unramified.* □

Proof of Theorem 3.1.1 – Let L be a field as in the proposition. Then we can identify $\Sigma' = \mathrm{Gal}(KL/L)$ with $\Sigma = Gal(K/k)$. Let ν be any place in k which does not ramify in K; then it does not ramify in L nor in KL. Denoting by w a place in KL over ν and the intermediate places in K and in L, we can also identify the local Galois groups $\mathrm{Gal}((KL)_w/L_w)$ and $\mathrm{Gal}(K/w)$.

We have now $H^1_{\mathrm{loc}}(\Sigma', A) \neq 0$, but this time Z vanishes in $H^1(\Sigma', \mathcal{A}(LK))$. As in the above discussion, we conclude the existence of an L-rational point P on $\mathcal{A}$ as required. □

Now, assume that Theorem 1.3.4 holds for some prime $p \neq 2$ and some positive integer n. As we have recalled in Chapter 1, for any normal field extension K/k and any surjective homomorphism $\psi\colon \mathrm{Gal}(K/k) \to \Sigma$ we can construct an algebraic torus $\mathcal{T}$ of dimension n, defined over k. By the arguments in Chapter 2, this allows us to read the condition $H^1_{\mathrm{loc}}(G, \mathbb{F}_p^n) \neq 0$ as $H^1_{\mathrm{loc}}(\Sigma, A) \neq 0$. Possibly enlarging the base field k, we obtain a counterexample to the local-global divisibility problem $(p, k, \mathcal{T})$. This proves Theorem 1.3.2. (Note that in the theorem we are not interested in the number field k, but only in the dimension n of $\mathcal{T}$, with respect to p.)

3.2 Proof of Theorem 1.3.4

In this section we prove Theorem 1.3.4, completing the proof of Theorem 1.3.2. Thus, let $p \neq 2$ be a prime and let $n \geqslant 3(p-1)$ be an integer. We will construct a p-group G of matrices in $\mathrm{SL}_n(\mathbb{Z})$ and a $(G, \mathbb{F}_p^n)$-cocycle Z that satisfies the local conditions without being a coboundary.

We recall that any example in the case $n = 3(p-1)$ can be extended to the general case by means of a direct sum with a trivial representation of dimension $n - 3(p-1)$. We can thus assume $n = 3(p-1)$.

Let $M \in \mathrm{SL}_{p-1}(\mathbb{Z})$ be any matrix with minimal polynomial $(x^p - 1)/(x - 1)$, for instance, the Frobenius matrix of this polynomial. Note that M satisfies Proposition 2.2.2. Let now $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{Z}^{p-1}$ such that

$$\mathbf{u} \not\equiv O \pmod{p}, \qquad \mathbf{v} \not\equiv O \pmod{p};$$
$$(M - I)\mathbf{u} \equiv O \pmod{p}, \qquad \mathbf{v}^t(M - I) \equiv O \pmod{p}.$$

We define the matrix $X := \frac{1}{p}\mathbf{u} \times \mathbf{v}^t$, with entries in $\mathbb{Q}$. Note that it is not an integer matrix. We also define the matrices $A := (M - I)X$ and $B := X(I - M)$, with entries in $\mathbb{Z}$.

Let G be the group generated by the matrices σ and τ defined as

$$\sigma = \begin{pmatrix} M & O & A \\ & M & A \\ & & I \end{pmatrix}, \qquad \tau = \begin{pmatrix} I & O & B \\ & M & A+B \\ & & M \end{pmatrix};$$

it is easily verified that G is a subgroup of $\mathrm{SL}_n(\mathbb{Z})$ and that the map

$$(i,j) \quad \mapsto \quad \sigma^i\tau^j = \begin{pmatrix} M^i & O & M^iX - XM^j \\ & M^{i+j} & M^{i+j}X - XM^j \\ & & M^j \end{pmatrix}$$

provides an isomorphism $G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

Lemma 3.2.1 *There exist vectors* $\mathbf{r}$, $\mathbf{s}$ *and* $\mathbf{t}$ *in* $\mathbb{Z}^{p-1}$ *such that:*

$$\begin{aligned} B\mathbf{t} &\equiv (M-I)\mathbf{r} \not\equiv O && \pmod p, \\ (M-I)B\mathbf{t} &\equiv O && \pmod p, \\ (A+B)\mathbf{t} &\equiv (M-I)\mathbf{s} && \pmod p. \end{aligned}$$

Proof – Assume $B(M-I)^{p-2} \not\equiv O \pmod p$. By Proposition 2.2.2 there exists an integer matrix X_0 with $B \equiv X_0(M-I) \pmod p$; since $(M-I)$ has determinant p, this implies that $X = B(M-I)^{-1}$ is an integer matrix. As X is not an integer matrix, this shows $B(M-I)^{p-2} \not\equiv O \pmod p$. In particular, there exists a vector $\mathbf{t}_0$ in $\mathbb{Z}^{p-1}$ such that $B(M-I)^{p-2}\mathbf{t}_0 \not\equiv O \pmod p$. Let $\mathbf{t} = (M-I)^{p-2}\mathbf{t}_0$, so that $B\mathbf{t} \not\equiv O \pmod p$.

By definition of A and B we have $(M-I)B = -A(M-I)$. Together with $(M-I)^{p-1} \equiv O \pmod p$, this implies $(M-I)B(M-I)^{p-2} \equiv O \pmod p$ and $(M-I)^{p-2}A(M-I) \equiv O \pmod p$. It follows that $(M-I)B\mathbf{t} \equiv O \pmod p$ and $(M-I)^{p-2}(A+B)\mathbf{t} \equiv O \pmod p$; we conclude by Proposition 2.2.2. □

Proposition 3.2.2 *The vectors* $Z^{(1)}_\sigma = O$ *and* $Z^{(1)}_\tau = B\mathbf{t}$ *define a* $(G, \mathbb{F}_p^n)$*-cocycle* $Z \equiv \begin{pmatrix} Z^{(1)} \\ O \\ O \end{pmatrix} \pmod p$ *that is not a* $(G, \mathbb{F}_p^n)$*-coboundary.*

Proof – To show that Z is a cocycle we only need to verify, on $Z^{(1)}$, the cocycle conditions derived from the relations $\sigma^p = I$, $\tau^p = I$ and $\sigma\tau = \tau\sigma$:

$$\begin{aligned} Z^{(1)}_{\sigma^p} - Z^{(1)}_I &\equiv (M^{p-1} + \ldots + M + I)Z^{(1)}_\sigma \equiv O && \pmod p; \\ Z^{(1)}_{\tau^p} - Z^{(1)}_I &\equiv pZ^{(1)}_\tau \equiv O && \pmod p; \\ Z^{(1)}_{\sigma\tau} - Z^{(1)}_{\tau\sigma} &\equiv (M-I)Z^{(1)}_\tau \equiv O && \pmod p. \end{aligned}$$

If Z was a coboundary, then there would exist a vector W in $\mathbb{Z}^n$ such that $Z_g \equiv (g-I)W \pmod p$ for every g in G; computing Z_σ and Z_τ, we would obtain

$$\begin{aligned} Z^{(2)}_\sigma &\equiv (M-I)W^{(2)} + AW^{(3)} && \pmod p, \\ Z^{(1)}_\tau &\equiv BW^{(3)} && \pmod p, \\ Z^{(2)}_\tau &\equiv (M-I)W^{(2)} + AW^{(3)} + BW^{(3)} && \pmod p, \end{aligned}$$

which is absurd, since $Z_\tau^{(2)} \equiv Z_\sigma^{(2)} \equiv O \pmod p$ and $Z_\tau^{(1)} \not\equiv O \pmod p$. □

We can now prove our theorem.

Proof of Theorem 1.3.4 – Let G and Z be as above. Then $[Z]$ is a non-trivial element of $H^1(G, \mathbb{F}_p^n)$. We can prove $H^1_{\mathrm{loc}}(G, \mathbb{F}_p^n) \neq 0$ by showing that $[Z]$ belongs to $H^1_{\mathrm{loc}}(G, \mathbb{F}_p^n)$, i.e. that for every g in G there exists a W_g in $\mathbb{F}_p^n$ such that $Z_g \equiv (g - I)W_g \pmod p$.

Over τ we have

$$(\tau - I)\begin{pmatrix} O \\ -\mathbf{s} \\ \mathbf{t} \end{pmatrix} \equiv \begin{pmatrix} O & O & B \\ O & M-I & A+B \\ O & O & M-I \end{pmatrix}\begin{pmatrix} O \\ -\mathbf{s} \\ \mathbf{t} \end{pmatrix} \equiv \begin{pmatrix} Z_\tau^{(1)} \\ O \\ O \end{pmatrix} \pmod p$$

For every $i \in \mathbb{F}_p^*$ we have $Z^{(1)}_{\tau^i\sigma} \equiv iZ_\tau^{(1)} + Z_\sigma^{(1)} \equiv iB\mathbf{t} \pmod p$; then

$$(\sigma\tau^i - I)\begin{pmatrix} i\mathbf{r} \\ O \\ O \end{pmatrix} \equiv \begin{pmatrix} M-I & \star & \star \\ O & \star & \star \\ O & O & \star \end{pmatrix}\begin{pmatrix} i\mathbf{r} \\ O \\ O \end{pmatrix} \equiv \begin{pmatrix} Z^{(1)}_{\sigma\tau^i} \\ O \\ O \end{pmatrix} \pmod p$$

Since τ and the $\sigma\tau^i$ with $i \in \mathbb{F}_p$ are the generators of all non-trivial cyclic subgroups of G, this shows that Z satisfies the local conditions. This completes the proof. □

Part II

Effective Diophantine Analysis on Modular Curves

Chapter 4

Introduction

4.1 Modular curves

We briefly recall a few definitions and notations concerning congruence subgroups and modular curves. For all missing details one may consult, for instance, [16].

Every matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ acts on the projective space $\mathbb{P}^1(\mathbb{C})$ as $M(z) = \frac{az+b}{cz+d}$. The two matrices $\pm M$ have the same action, so that, actually, the action is given by matrices in $\mathrm{PSL}_2(\mathbb{Z})$. We have the stable sets $\mathbb{Q} \cup \{\infty\} \subset \mathbb{P}^1(\mathbb{R})$ and $\mathcal{H} = \{z = x + yi \mid y > 0\}$.

Every finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ acts on

$$\bar{\mathcal{H}} = \mathcal{H} \cup \mathbb{Q} \cup \{\infty\};$$

the quotient set $\Gamma \backslash \bar{\mathcal{H}}$, supplied with the properly defined topology and analytic structure, give a Riemann surface X_Γ; by the Riemann existence theorem, X_Γ is a complex algebraic curve, known as *modular curve*. We shall usually assume that Γ contains the matrix $-I$; this will enable us to consider it as the pull-back of $\bar{\Gamma} \leq \mathrm{PSL}_2(\mathbb{Z})$.

There are some algebraic invariants of Γ (or X_Γ), i.e. invariants by conjugation in $\mathrm{PSL}_2(\mathbb{Z})$. Let us see them.

- The *genus* $\mathbf{g}$ of the modular curve X_Γ.
- The *index* $\mu = [\mathrm{PSL}_2(\mathbb{Z}) \colon \Gamma]$.
- The number of *cusps*, $\nu_\infty = |\Gamma \backslash (\mathbb{Q} \cup \{\infty\})|$.
- The numbers ν_2 of 2-*elliptic points* and ν_3 of 3-*elliptic points*.

A non-cuspidal point $P \in X_\Gamma$ is called *elliptic* if for some $z \in \mathcal{H}$ representing P (or, equivalently, for any such z) the stabilizer $\bar{\Gamma}_z$ (which is always finite) is non-trivial. It is known that every non-trivial finite subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ is cyclic of order 2 or 3. Hence the stabilizer can only be a cyclic group of order 2 or 3, which gives rise to 2-*elliptic points* or 3-*elliptic points*, respectively.

Remark 1 *A non-trivial finite order element of* $\mathrm{PSL}_2(\mathbb{Z})$ *is called elliptic; thus, we have 2-elliptic and 3-elliptic elements. All the 2-elliptic elements are* $\mathrm{PSL}_2(\mathbb{Z})$*-conjugate; for the 3-elliptic elements, there are 2 conjugacy classes; if* M *is 3-elliptic, then, one class is that of* M *and the other that of* M^2*.*

An element of $\mathrm{SL}_2(\mathbb{Z})$ *is 2- (respectively, 3-) elliptic if its image in* $\mathrm{PSL}_2(\mathbb{Z})$ *is 2- (respectively, 3-) elliptic. The 2-elliptic elements of* $\mathrm{SL}_2(\mathbb{Z})$ *are of order 4 and trace 0. The 3-elliptic elements of* $\mathrm{SL}_2(\mathbb{Z})$ *are of order 3 or 6 and trace* -1 *or 1, respectively.*

The invariants defined above are connected by the celebrated *Hurwitz formula*:

$$\mathrm{g} = 1 + \frac{\mu}{12} - \frac{\nu_\infty}{2} - \frac{\nu_3}{3} - \frac{\nu_2}{4}. \tag{4.1}$$

The classical j-invariant function, defined on $\mathcal{H}$ by the familiar relation

$$j(z) = q^{-1} + 744 + 196844q + \ldots,$$

where $q = e^{2\pi iz}$, is $\mathrm{SL}_2(\mathbb{Z})$-automorphic; hence it is Γ-automorphic for any $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$. It follows that j defines a function on X_Γ, which can be shown to be meromorphic, with poles exactly at the cusps. We have $j(P) = 0$ if P is a 2-elliptic point, and $j(P) = 1728$ if P is a 3-elliptic point.

The field $\mathbb{C}(j)$ is a subfield of $\mathbb{C}(X_\Gamma)$ of index $\mu = \mu(\Gamma)$. In other words, j defines a covering $X_\Gamma \to \mathbb{P}^1$ of degree μ. This covering is unramified over $\mathbb{P}^1 \setminus \{0, 1728, \infty\}$.

So far we defined the modular curve X_Γ as a complex algebraic curve, and j as an element of $\mathbb{C}(X_\Gamma)$. It turns out that, for any finite index subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ the curve X_Γ is definable over some number field k, in the way that j becomes an element of $k(X_\Gamma)$. This is the "easy" part of the celebrated theorem of Belyi [15, page 71], which asserts that for a complex algebraic curve X the following three conditions are equivalent:

- X is $\mathbb{C}$-isomorphic to X_Γ for some $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$;
- X admits a finite covering of $\mathbb{P}^1$, unramified outside the points 0, 1 and ∞;
- X can be defined over a number field.

Finally, remark that the curve X_Γ can be *defined effectively* in the following sense. There exists a number field k, and a polynomial $f(T, Y) \in k[T, Y]$ such that the following holds.

- The degree and discriminant of k, as well as the degree and the height[1] of f are bounded effectively in terms of $\mu(\Gamma)$.
- The curve X_Γ is definable over k and $j \in k(X_\Gamma)$.
- There exists $y \in k(X_\Gamma)$ such that $k(X_\Gamma) = k(j, y)$ and the functions j and y satisfy the polynomial equation $f(j, y) = 0$.

This is well-known and may be viewed as a particular case of the "effective Riemann existence theorem", as in [4, Chapter 3]. A totally explicit version of this statement will appear in the forthcoming thesis of M. Strambi [17].

[1] By the height of a (non-zero) polynomial we mean the projective Weil height of the vector of its coefficients.

Congruence subgroups

Everything above is true for any finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$. However, in this thesis we mainly deal with the congruence subgroups. Recall that the principal congruence subgroup $\Gamma(n)$ of $\mathrm{SL}_2(\mathbb{Z})$ is the kernel of the reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/(n))$. We say that a subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a *congruence subgroup of level* n if Γ contains $\Gamma(n)$. The minimal n with this property will be called the *exact level* of Γ; it divides every other level of Γ.

For every positive integer n, there are two classical congruence subgroups of exact level n:

$$\Gamma_1(n) = \left\{ M \in \mathrm{SL}_2(\mathbb{Z}) \mid M \equiv \begin{pmatrix} 1 & \star \\ 0 & 1 \end{pmatrix} \pmod{n} \right\};$$
$$\Gamma_0(n) = \left\{ M \in \mathrm{SL}_2(\mathbb{Z}) \mid M \equiv \begin{pmatrix} \star & \star \\ 0 & \star \end{pmatrix} \pmod{n} \right\}.$$

The modular curves corresponding to the groups $\Gamma(n)$, $\Gamma_1(n)$, and $\Gamma_0(n)$ are usually denoted by $X(n)$, $X_1(n)$, and $X_0(n)$ respectively.

For congruence subgroups, the above mentioned definability of the modular curve over a number field can be made very explicit. For instance, the curve $X(n)$ can be defined over the cyclotomic field $\mathbb{Q}(\zeta_n)$, and the same is true for any X_Γ with Γ of level n. In some special cases even more can be said: for instance, $X_0(n)$ can be defined over $\mathbb{Q}$. (See, for instance, [12].)

Similarly, in the congruence case one can give explicit defining equations for the corresponding modular curves ("modular equations"). See [12] for such an equation for the curve $X_0(n)$.

4.2 The theorem of Siegel

To describe our problem, we recall the classical theorem of Siegel. Let X be a projective curve defined over a number field k and j a non-constant element of $k(X)$ (a "coordinate"). Further, let K be a finite extension of k, let S be a finite set of places on K (which includes all the infinite places), and $\mathcal{O}_S$ the ring of S-integers of the field K. We define the set of S-integral points on X with respect to the coordinate j as follows:

$$X(\mathcal{O}_S, j) = \{P \in X(K) \mid j(P) \in \mathcal{O}_S\}.$$

Theorem 4.2.1 (Siegel) *Assume that either* $\mathbf{g}(X) \geq 1$ *or* j *has at least* 3 *poles. Then for any* K *and* S *as above, the set* $X(\mathcal{O}_S, j)$ *is finite.*

For a modern proof of this theorem, one may consult [15].

A pair (X, j) satisfying the assumption of this theorem (that is, *either* $\mathbf{g}(X) \geq 1$ *or* j *has at least* 3 *poles*) will be called *Siegelian*. Thus, Siegel's theorem asserts that for a Siegelian pair (X, j), the set of S-integral points on X with respect to j is finite.

Remark that the converse statement is also true: if the set $X(\mathcal{O}_S, j)$ is finite *for all* K and S as above, then the pair (X, j) is Siegelian. For a non-Siegelian pair, the set $X(\mathcal{O}_S, j)$ can be finite or infinite; see [1] for a finiteness criterion for non-Siegelian pairs.

Siegel's theorem states that the set of integral points on X is finite, but, unfortunately, its proof does not imply any upper bound on the size of integral points. (By the *size* of a point $P \in X(\mathcal{O}_S, j)$ we mean the height of the algebraic number $j(P)$.)

Starting from pioneering work of A. Baker, there have been obtained effective versions for some cases of this theorem; see [5, 6] for the history of the subject and further references. For instance, the following is known.

Theorem 4.2.2 *Siegel's theorem is effective for (X, j) if*

1. (folklore) $\mathbf{g}(X) = 0$ *and j has at least 3 poles, or*
2. (Baker and Coates [2]) $\mathbf{g}(X) = 1$, *or*
3. (Bilu [3], Dvornicich and Zannier [8]) $\mathbf{g}(X) \geq 1$ *and $\bar{k}(X)/\bar{k}(j)$ is a Galois extension.*

We say that *Siegel's theorem is effective* for the Siegelian pair (X, j) if for any K and S as above, the sizes of points from $X(\mathcal{O}_S, j)$ are effectively bounded in terms of K, S and (X, j). Here *effectively bounded in terms of K and S* means bounded by a constant explicitly depending on the degree and discriminant of K, and the maximal (finite) prime number below S. Further, effectively bounded in terms of (X, j) means bounded by a constant which can be explicitly expressed in terms of the degree and the height of the defining equation of some plane model of X, such that j is one of the coordinates of this model. (As we have seen in the previous section, for the pair (X_Γ, j) such a model can be effectively determined.)

For the modular curves the following effective results have been established in [5, 6].

Theorem 4.2.3 *Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and j be the j-invariant. Siegel's theorem is effective for the pair (X_Γ, j) if one of the following conditions is satisfied.*

1. *The group Γ is a congruence subgroup and X_Γ has at least 3 cusps.*
2. *The group Γ has no elliptic elements.*

Bilu [5, 6] also proved effective Siegel's theorem for (X_Γ, j) when Γ is one of the classical congruence subgroups $\Gamma(N)$, $\Gamma_1(N)$ and $\Gamma_0(N)$, provided the corresponding pair is Siegelian, that is (see [6, Corollary 9 and Theorem 10])

- for $(X(n), j)$, when $n \geqslant 2$;
- for $(X_1(n), j)$, when $n \geqslant 4$;
- for $(X_0(n), j)$, when $n \notin \{1, 2, 3, 5, 7, 13\}$.

4.3 Main results

In this thesis we establish effective Siegel's theorem for (X_Γ, j) where Γ is "almost every" congruence subgroup. In the prime power level our result is nearly best possible (see Chapters 6 and 7).

Theorem 4.3.1 *Let Γ be a congruence subgroup of the exact prime power level p^n. Assume that $p^n \neq 25$. Then either Siegel's theorem is effective for the pair (X_Γ, j), or this pair is non-Siegelian.*

Actually, we obtain much more precise statements. In particular, we explicitly exhibit the "nasty" subgroup of level 25 for which our method does not work. Also, we classify all Γ of the prime power level for which the pair (X_Γ, j) is not Siegelian.

In the general case we prove the following (see Chapter 8).

Theorem 4.3.2 *Let Γ be a congruence subgroup of (exact) level not dividing $2^{21} \cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 11 \cdot 13$. Then Siegel's theorem is effective for the pair (X_Γ, j).*

Again, more precise statements are available.

Chapter 5

Counting cusps and elliptic points

This chapter is of technical nature. We describe here some tools for enumerating cusps and elliptic points of a modular curve.

Let Γ be a congruence subgroup of level dividing n, containing $-I$. We consider its projection modulo n, the group $G_n < \mathrm{SL}_2(\mathbb{Z}/(n))$ isomorphic to $\Gamma/\Gamma(n)$, and its image $\bar{G}_n = G/\{\pm I\}$ in $\mathrm{PSL}_2(\mathbb{Z}/(n))$. For n fixed we shall write $G = G_n$ and $\bar{G} = \bar{G}_n$.

In this section we show how the numbers of cusps and of elliptic points, the index, and therefore the genus of the modular curve X_Γ can be obtained via the finite groups G and $\bar{G}$.

For instance, the index $\mu(\Gamma)$ is

$$\mu(\Gamma) = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma] = [\mathrm{SL}_2(\mathbb{Z}/(n)) : G] = [\mathrm{PSL}_2(\mathbb{Z}/(n)) : \bar{G}].$$

Most or even all of the statement in this section are certainly well-known, but we include the proofs for completeness.

5.1 Enumerating the cusps

Let $\mathcal{M}_n$ be the set of elements of the exact order n (i.e. of the maximal possible order) in the abelian group $\mathbb{Z}/(n) \times \mathbb{Z}/(n)$. In this section we shall prove the following result.

Theorem 5.1.1 *Let Γ be a congruence subgroup of level dividing n and containing $-I$, and let G be the projection of Γ modulo n. Then the number $\nu_\infty(\Gamma)$ equals the number of G-orbits of $\mathcal{M}_n$. In other words, we have $\nu_\infty(\Gamma) = \#(G\backslash\mathcal{M}_n)$.*

We recall that, by definition, $\nu_\infty(\Gamma)$ is the number of Γ-orbits of the set $\mathbb{Q} \cup \{\infty\} = \mathbb{P}^1(\mathbb{Q})$. Let $\mathcal{M}$ be the set of couples of coprime integers (x, y), and note that the projection modulo n maps $\mathcal{M}$ onto $\mathcal{M}_n$.

Proposition 5.1.2 *Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ containing $-I$. Then the map $f\colon \mathcal{M} \to \mathbb{P}^1(\mathbb{Q})$ given by $f(x, y) = \frac{x}{y}$ defines a bijection between $\Gamma\backslash\mathcal{M}$ and $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$.*

Proof – The action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{Q})$ is induced by the natural action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{Z} \times \mathbb{Z}$, and f is a 2-to-1 surjective map. In particular, f defines an $\mathrm{SL}_2(\mathbb{Z})$-equivariant bijection between $\{\pm I\}\backslash\mathcal{M}$ and $\mathbb{P}^1(\mathbb{Q})$. Whence the result. □

Proposition 5.1.3 *Let $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathcal{M}$ be a pull-back of $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \in \mathcal{M}_n$. Then there exists an element $N \in \Gamma(n)$ such that $N\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$.*

Proof – Since x and y are coprime, there exist integers u and t such that $xu - yt = 1$. The matrix $N = \left(\begin{smallmatrix} x & t-xt \\ y & u-yt \end{smallmatrix}\right)$ lies in $\Gamma(n)$ and satisfies the required relation. □

Proposition 5.1.4 *Let Γ be a congruence subgroup of level dividing n, and let $G \cong \Gamma/\Gamma(n)$ be its image modulo n. Then the projection modulo n defines a bijection between $\Gamma\backslash\mathcal{M}$ and $G\backslash\mathcal{M}_n$.*

Proof – The projection map $\pi_n \colon \mathcal{M} \to \mathcal{M}_n$ induces an action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{M}_n$, which factors through $\Gamma/\Gamma(n) \cong G$. By Proposition 5.1.3, the group $\Gamma(n)$ acts transitively on the fiber $\pi_n^{-1}\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$; since $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathcal{M}$, and since $\Gamma(n)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$, then $\Gamma(n)$ acts transitively on the fiber of every element of M_n. Thus π_n defines an $\mathrm{SL}_2(\mathbb{Z})$-equivariant bijection between $\Gamma(n)\backslash\mathcal{M}$ and $\mathcal{M}_n$. Whence the result. □

Proof of Theorem 5.1.1 – Immediately from Propositions 5.1.2 and 5.1.4. □

Recall that we are interested in the congruence groups Γ with at most 2 cusps. The following propositions will give us some helpful information.

Proposition 5.1.5 *Let Γ be a congruence subgroup of level n containing $-I$ and let G be its image in $\mathrm{SL}_2(\mathbb{Z}/(n))$. Assume that Γ has at most 2 cusps. Then $\#G \geqslant \#\mathcal{M}_n/2$.*

Proof – This is an immediate consequence of Theorem 5.1.1. □

Proposition 5.1.6 *The set $\mathcal{M}_n$ has cardinality*

$$\#\mathcal{M}_n = n^2 \prod_{p|n} \left(1 - p^{-2}\right),$$

the product being taken over all primes p that divide n.

Proof – By the Chinese Remainder Theorem, the function $f(n) = \#\mathcal{M}_n$ is multiplicative, and for $n = p^k$ the statement is obvious. □

5.2 Enumerating the elliptic points

Let $\mathcal{E}_2$ and $\mathcal{E}_3$ be the sets of 2- and 3-elliptic elements of $\mathrm{PSL}_2(\mathbb{Z})$, and let $\mathcal{E}_2(n)$ and $\mathcal{E}_3(n)$ be their projections in $\mathrm{PSL}_2(\mathbb{Z}/(n))$. In this section we shall prove the following result.

Theorem 5.2.1 *Let Γ be a congruence subgroup of level dividing n and containing $-I$, and let $\bar{G} \cong \Gamma/\langle -I, \Gamma(n)\rangle$ be its image in $\mathrm{PSL}_2(\mathbb{Z}/(n))$. There exist functions f_2 and f_3, depending only on n, such that*

$$\nu_2(\Gamma) = \frac{\#(\bar{G} \cap \mathcal{E}_2(n))}{\#\bar{G}} f_2(n); \qquad \nu_3(\Gamma) = \frac{\#(\bar{G} \cap \mathcal{E}_3(n))}{\#\bar{G}} f_3(n).$$

Let $\bar{\Gamma} \cong \Gamma/\{\pm I\}$ be the image of Γ in $\mathrm{PSL}_2(\mathbb{Z})$. We recall that, by definition, $\nu_2(\Gamma)$ and $\nu_3(\Gamma)$ are the numbers of distinct $\bar{\Gamma}$-conjugacy classes of elements of exact order 2 and 3 of $\bar{\Gamma}$; the elements of exact order 2 are the traceless elements, and the elements of order 3 are those of trace ± 1.

Let H be a group acting on a set X. We denote by x^H the H-orbit of any element x of X. Let also N be a normal subgroup of H; we denote by a tilde the quotient modulo N.

Proposition 5.2.2 *Let H be a group and let N be a normal subgroup of finite index of H. Consider the action of H on itself by conjugacy. If the centralizer in H of an element x is $Z_H(x) = \langle x\rangle$, then the conjugacy class of x in H is union of $\#\widetilde{H}/ord(\widetilde{x})$ many N-conjugacy classes.*

Proof – The conjugacy class x^H of x in H is a finite union $x_1^N \cup \ldots \cup x_r^N$ of distinct N-conjugacy classes. The action of H on x^H induces an action on the set $\{x_1^N, \ldots, x_r^N\}$, which factors thorugh $\widetilde{H}$. Since any x_i is conjugate to x, its stabilizer in $\widetilde{H}$ is $\langle \widetilde{x}_i\rangle$. □

Proposition 5.2.3 *Let H be a group, let N be a normal subgroup of finite index of H, and let X be a non-empty union of H-conjugacy classes of elements whose images in $\widetilde{H}$ have the same exact order r. For every subgroup K of H containing N, we denote by n_K the number of distinct K-conjugacy classes in $X_K = X \cap K$. If, for every x in X, the centralizer of x in H is $\langle x\rangle$, then the number*

$$n_K \frac{\#\tilde{K}}{\#\tilde{X}_K}$$

does not depend on K.

Proof – By the above proposition, the numbers of N-orbits in X and in X_K are respectively $n_H \#\widetilde{H}/k$ and $n_K \#\widetilde{K}/k$. By the same proposition, for every x in X, the number of distinct N-orbits in $X \cap xN$ does not depend on x. This implies that the ratio among the numbers of N-orbits in X_K and in X is $\#\widetilde{X}_K/\widetilde{X}$. Thus we have

$$\frac{\#\widetilde{X}_K}{\#\widetilde{X}} = \frac{n_K \#\widetilde{K}}{n_H \#\widetilde{H}},$$

which completes the proof. □

Proof of Theorem 5.2.1 – In the above proposition, let $H = \mathrm{PSL}_2(\mathbb{Z})$, $K = \bar{\Gamma}$, and $N = \bar{\Gamma}(n) = \langle \bar{\Gamma}, -I \rangle / \{\pm I\}$. Taking either $X = \mathcal{E}_2$ or $X = \mathcal{E}_3$, we obtain that the functions

$$f_2(n) = \nu_2(\Gamma) \frac{\#\bar{G}}{\#(\bar{G} \cap \mathcal{E}_2(n))},$$
$$f_3(n) = \nu_3(\Gamma) \frac{\#\bar{G}}{\#(\bar{G} \cap \mathcal{E}_3(n))}$$

depend only on n. □

We recall that the elements of $\mathcal{E}_2 \subset \mathrm{PSL}_2(\mathbb{Z})$ are characterized by their order, or by their trace, and so are the elements of $\mathcal{E}_3 \subset \mathrm{PSL}_2(\mathbb{Z})$. Considering conjugacy classes in $\mathrm{SL}_2(\mathbb{Z}/(n))$, we obtain the following partial characterization of the sets $\mathcal{E}_2(n)$ and $\mathcal{E}_3(n)$ when n is a prime power.

Proposition 5.2.4 *Let $n = p^e$ is a prime power and let M be an element of $\mathrm{SL}_2(\mathbb{Z}/(n))$. For $p \neq 2$ the following properties are equivalent:*

$$M \in \mathcal{E}_2(n); \qquad \mathrm{ord}(M) = 2; \qquad \mathrm{Tr}(M) = 0.$$

For $p \neq 3$ the following properties are equivalent:

$$M \in \mathcal{E}_3(n); \qquad \mathrm{ord}(M) = 3; \qquad |\mathrm{Tr}(M)| = 1.$$

□

Taking $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ in Theorem 5.2.1, we can compute explicit values of f_2 and f_3.

Proposition 5.2.5 *The functions $f_2(n)$ and $2f_3(n)$ are multiplicative on n. For any prime p we have:*

$$f_2(p^e) = \begin{cases} p^e & \text{if } p = 2, \\ p^e(1 - p^{-1}) & \text{if } p \equiv 1 \pmod 4, \\ p^e(1 + p^{-1}) & \text{if } p \equiv -1 \pmod 4; \end{cases}$$
$$2f_3(p^e) = \begin{cases} p^e & \text{if } p = 3, \\ p^e(1 - p^{-1}) & \text{if } p \equiv 1 \pmod 3, \\ p^e(1 + p^{-1}) & \text{if } p \equiv -1 \pmod 3. \end{cases}$$

□

Chapter 6

The prime level case

6.1 Special groups

In this chapter we prove the following result.

Theorem 6.1.1 *Let Γ be a congruence subgroup of the (exact) level equal to 1 or to a prime number. Then Siegel's theorem is effective for the pair (X_Γ, j) whenever this pair is Siegelian.*

All the Γ (up to conjugacy) of level 1 or prime, for which the pair (X_Γ, j) is non-Siegelian, are listed in Tables 6.1 and 6.2.

Here we collect basic properties of the special linear group $\mathrm{SL}_2(\mathbb{F}_p)$.

The following property is well-known but we sketch a proof for the sake of completeness.

Proposition 6.1.2 *The order of an element of $\mathrm{SL}_2(\mathbb{F}_p)$ is either $2p$ or at most $p+1$. When $p \neq 2$, the order of an element of $\mathrm{PSL}_2(\mathbb{F}_p)$ is either p or at most $(p+1)/2$.*

Proof – A matrix from $\mathrm{SL}_2(\mathbb{F}_p)$ is either similar over $\mathbb{F}_p$ to $\left(\begin{smallmatrix}\lambda & 1\\ 0 & \lambda\end{smallmatrix}\right)$ with $\lambda = \pm 1$ or similar over $\mathbb{F}_{p^2}$ to $\left(\begin{smallmatrix}\alpha & 0\\ 0 & \alpha^{-1}\end{smallmatrix}\right)$ with $\alpha \in \mathbb{F}_{p^2}$. In the first case the order divides $2p$. In the second case either $\alpha \in \mathbb{F}_p$, in which case the order divides $p-1$, or α is in the kernel of the norm map $\mathbb{F}_{p^2} \to \mathbb{F}_p$, in which case the order divides $p+1$. □

We shall systematically use the classification of semi-simple subgroups of $\mathrm{PSL}_2(\mathbb{F}_p)$. Actually, a classification of $\mathrm{PGL}_2(\mathbb{F}_p)$ is available, see [13, Proposition 16].

Proposition 6.1.3 *Let $\bar{G}$ be a proper subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$ of order not divisible by p. Then $\bar{G}$ is isomorphic to one of the following groups:*

- *$\mathcal{C}_n$, the n-th cyclic group;*
- *$\mathcal{D}_n$, the n-th dihedral group;*
- *$\mathcal{A}_4$, the fourth alternated group;*
- *$\mathcal{S}_4$, the fourth symmetric group;*

- $\mathcal{A}_5$, *the fifth alternated group (this only happens when* $p \equiv \pm 1 \pmod 5$*).*

In the unipotent case, one has the following, see [13, Proposition 15].

Proposition 6.1.4 *Let G be a subgroup of $GL_2(\mathbb{F}_p)$ of order divisible by p. Then G either contains $\mathrm{SL}_2(\mathbb{F}_p)$ or is contained in a Borel subgroup of $GL_2(\mathbb{F}_p)$.*

(A Borel subgroup of $\mathrm{GL}_2(\mathbb{F}_p)$ is any conjugate of the subgroup $\mathrm{GT}_2(\mathbb{F}_p)$ of the upper-triangular matrices.)

Proposition 6.1.5 *Let H be a subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$, conjugate to $\mathrm{ST}_2(\mathbb{F}_p)$, and let G be a subgroup of H, with $\nu_\infty(G) \leqslant 2$. Then $G = H$.*

Proof – If G were a proper subgroup of H, then its cardinality would be at most half the cardinality of H. Then

$$p^2 - p = \#\mathrm{ST}_2(\mathbb{F}_p) = \#H \geqslant 2\#G \geqslant \nu_\infty(G)\#G \geqslant \#\mathcal{M}_p = p^2 - 1,$$

which is absurd. □

Theorem 6.1.6 *Let Γ be a congruence subgroup of exact level p, with at most 2 cusps.*

- *If p does not divide the cardinality of $\bar{G}$ then we are in one of the following 8 cases.*
  - *$p = 2$ and $\bar{G} \cong \mathcal{C}_3 \cong \mathbb{Z}/(3)$;*
  - *$p = 3$ and $\bar{G} \cong \mathcal{C}_2 \cong \mathbb{Z}/(2)$;*
  - *$p = 3$ and $\bar{G} \cong \mathcal{D}_2 \cong \mathbb{Z}/(2) \times \mathbb{Z}/(2)$;*
  - *$p = 5$ and $\bar{G} \cong \mathcal{D}_3$;*
  - *$p = 5$ and $\bar{G} \cong \mathcal{A}_4$;*
  - *$p = 7$ and $\bar{G} \cong \mathcal{A}_4$;*
  - *$p = 7$ and $\bar{G} \cong \mathcal{S}_4$;*
  - *$p = 11$ and $\bar{G} \cong \mathcal{A}_5$;*
- *If p divides the cardinality of $\bar{G}$ then G is conjugate to $\mathrm{ST}_2(\mathbb{F}_p)$ and $\nu_\infty(\Gamma) = 2$.*

Proof – When p does not divide the cardinality of G, we apply Proposition 6.1.3 and Proposition 6.1.2, which provide a bound for the cardinality of G; Proposition 5.1.5 does the rest. The cases follow.

First, the case $p = 2$. We have $\mathrm{PSL}_2(\mathbb{F}_2) \cong \mathrm{SL}_2(F_2) \cong \mathcal{S}_3$. Its proper subgroups are cyclic, with order 1, 2, or 3. We exclude the index 1 by counting the orbits, and the index 3 by divisibility.

We now deal with the other cases, applying the formula $\nu_\infty(\Gamma)\#G \geqslant p^2 - 1$.

When $\bar{G} \cong \mathcal{C}_n$, we have $\#G = 2n \leqslant p + 1$. Then we obtain $p \leqslant 3$, i.e. $p = 3$.

When $\bar{G} \cong \mathcal{D}_n$, we have $\#G = 4n \leqslant 2(p + 1)$. This gives $p \leqslant 5$, i.e. $p = 3, 5$. From $n = \frac{1}{4}\#G \geqslant \frac{1}{8}(p^2 - 1)$ we obtain that $p = 3$ implies $n = 1, 2$ and p^5 implies $n = 3$.

When $\bar{G} \cong \mathcal{A}_4$, we have $p^2 - 1 \leqslant 4\#\bar{G} = 48$, i.e. $p \leqslant 7$. Since 3 divides $\#\bar{G}$, we obtain $p = 5, 7$.

When $\bar{G} \cong \mathcal{S}_4$, we have $p^2 - 1 \leqslant 4\#\bar{G} = 96$, i.e. $p < 10$. Since $\#\mathcal{S}_4 = 24$ does not divide neither $\#\mathrm{PSL}_2(\mathbb{F}_3) = 12$ nor $\#\mathrm{PSL}_2(\mathbb{F}_5) = 60$, we have $p = 7$.

Finally, when $\bar{G} \cong \mathcal{A}_5$ we have $p \equiv \pm 1 \pmod 5$ and $p^2 - 1 \leqslant 4\#\bar{G} = 240$, i.e. $p = 11$.

When p divides the order of G, by Proposition 6.1.4 either $G = \mathrm{SL}_2(F_p)$ or G is conjugate to a subgroup of $\mathrm{ST}_2(\mathbb{F}_p)$. In the first case we have that Γ is the whole $\mathrm{SL}_2(\mathbb{Z})$, against our assumption on its level. In the second case we apply Proposition 6.1.5. □

Consider the 8 cases where p does not divide the order of $\bar{G}$. We remark that in the first five cases (with $p \leqslant 5$) the group Γ is uniquely defined up to conjugacy, and that in each of the last three cases (with $p \geqslant 7$) the group Γ can belong to two distinct conjugacy classes.

6.2 The 8 special cases

In this section we shall show that the 8 cases when p does not divide the cardinality of G listed in Theorem 6.1.6 all have genus $\mathbf{g} = 0$ and therefore do not satisfy the hypothesis of Siegel's theorem.

With the aid of Theorems 5.1.1 and 5.2.1, we obtain the invariants of the modular curve X_Γ by the isomorphism class of $\bar{G}$.

Table 6.1: Non-Siegelian modular curves of prime level: the semi-simple case

G	p	μ	ν_∞	ν_2	ν_3	$\mathbf{g}$	
$G \cong \mathcal{C}_3$	2	2	1	0	2	0	
$\bar{G} \cong \mathcal{C}_2$	3	6	2	2	0	0	
$\bar{G} \cong \mathcal{D}_2$	3	3	1	3	0	0	
$\bar{G} \cong \mathcal{D}_3$	5	10	2	2	1	0	
$\bar{G} \cong \mathcal{A}_4$	5	5	1	1	2	0	
$G \cong \mathcal{A}_4$	7	14	2	2	2	0	2 groups
$\bar{G} \cong \mathcal{S}_4$	7	7	1	3	1	0	2 groups
$\bar{G} \cong \mathcal{A}_5$	11	11	1	3	2	0	2 groups

Remark that in all the above cases we have $\nu_\infty(G)\#G = \#\mathcal{M}_p$.

In the first five lines of Table 6.1 the corresponding group Γ is well-defined up to $\mathrm{SL}_2(\mathbb{Z})$-conjugation. In every of the last three lines there are 2 conjugacy classes of Γ. Thus, Table 6.1 describes 11 possible Γ up to conjugacy.

6.3 The groups with order divisible by p

In this section we study the groups Γ for which p divides the cardinality of G and $\mu_\infty(\Gamma) \leq 2$. As we have seen in Theorem 6.1.6, in this case G is either conjugate to $\mathrm{ST}_2(\mathbb{F}_p)$ or is $\mathrm{SL}_2(\mathbb{F}_p)$ itself. In both cases we can assume, up to conjugacy, $\Gamma = \Gamma_0(n)$, for some positive integer n. (Note that $\Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$.)

A complete study on effectivity for this case is due to Bilu (see [6, Theorem 10]).

Theorem 6.3.1 *Either Siegel's theorem is effective for $(X_0(n), j)$ or the couple $(X_0(n), j)$ is non-Siegelian. The latter case is verified precisely when n is in the set $\{1, 2, 3, 5, 7, 13\}$.* □

Table 6.2: Non-Siegelian modular curves of level dividing a prime: the case $\Gamma = \Gamma_0(n)$

G	level	μ	ν_∞	ν_2	ν_3	g
$\mathrm{SL}_2(\mathbb{F}_p)$	1	1	1	1	1	0
$\mathrm{ST}_2(\mathbb{F}_2)$	2	3	2	1	0	0
$\mathrm{ST}_2(\mathbb{F}_3)$	3	4	2	0	1	0
$\mathrm{ST}_2(\mathbb{F}_5)$	5	6	2	2	0	0
$\mathrm{ST}_2(\mathbb{F}_7)$	7	8	2	0	2	0
$\mathrm{ST}_2(\mathbb{F}_{13})$	13	14	2	2	2	0

Chapter 7

The prime power level case

7.1 Introduction

In this chapter we study groups of prime power level. Our ultimate goal is the following theorem.

Theorem 7.1.1 *Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of prime power level, distinct from 25. Then either the pair (X_Γ, j) is non-Siegelian, or Siegel's theorem is effective for (X_Γ, j).*

In level 25 there is a subgroup Γ, defined below, for which the curve X_Γ is of genus 2 and for which our methods fail.

As in the prime case, our main tool will be "three cusps criterion" (Theorem 4.2.3) in the following refined form, see [6, Proposition 12].

Theorem 7.1.2 *Let Γ have a subgroup Γ' satisfying the following:*

- *Γ' is a congruence subgroup;*
- *Γ' contains all elliptic elements of Γ;*
- *$X_{\Gamma'}$ has at least 3 cusps.*

Then Siegel's theorem is effective for (X_Γ, j).

We obtain a complete classification, up to conjugacy, of the groups Γ of prime power level, containing $-I$, satisfying the following two conditions:

- Γ has at most two cusps, and
- Γ is generated by its elliptic elements.

We call Γ *unipotent* if its image in $\mathrm{SL}_2(\mathbb{Z}/(p))$ has an element of order p.

7.2 Projections

Let p be a prime, $q = p^e$ be a power of p, and Γ be a congruence subgroup of exact level q; this means that q is the smallest positive integer m such that Γ contains the kernel $\Gamma(m)$ of the projection $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/(m))$.

For the moment we shall deal with the general framework, but in due time we shall make the further assumptions:

$$\Gamma \text{ contains } -I \qquad \textbf{and} \qquad \text{the curve } X_\Gamma \text{ has at most two cusps.} \tag{7.1}$$

For every positive integer s the projection $\mathrm{SL}_2(\mathbb{Z}) \twoheadrightarrow \mathrm{SL}_2(\mathbb{Z}/(p^s))$ factors through $\mathrm{SL}_2(\mathbb{Z}/(p^{s+1}))$. This gives the chain of surjective projections

$$\Gamma \twoheadrightarrow G_e \twoheadrightarrow G_{e-1} \twoheadrightarrow \cdots \twoheadrightarrow G_2 \twoheadrightarrow G_1,$$

where G_s is a subgroup of $\mathrm{SL}_2(\mathbb{Z}/(p^s))$, for every s. Denoting by Γ_s the preimage in $\mathrm{SL}_2(\mathbb{Z})$ of G_s, i.e. $\Gamma_s := \Gamma \cdot \Gamma(p^s)$, we obtain the chain of inclusions

$$\Gamma = \Gamma_e \subsetneq \Gamma_{e-1} \subset \cdots \subset \Gamma_2 \subset \Gamma_1,$$

where every Γ_s is a congruence subgroup of level dividing p^s; in particular, Γ_1 is a congruence subgroup of level dividing p. Note that if Γ satisfies (7.1) then so does every Γ_s; in particular, Γ_1 belongs to the finite set of groups that we have determined in the previous chapter.

In this section we will show how to determine all the congruence subgroups Γ_{s+1} that project on a given Γ_s; in other words, all groups G_{s+1} that project on a fixed G_s. This will allow us, in the subsequent sections, to determine by induction in the exponent s all congruence subgroups of level dividing p^s, which satisfy (7.1).

7.2.1 The kernel

Let G_{s+1} be a subgroup of $\mathrm{SL}_2(\mathbb{Z}/(p^{s+1}))$, let π_s denote the projection modulo p^s, and let $G_s = \pi_s(G_{s+1})$ and $K_{s+1} = \mathrm{Ker}(\pi_s|_{G_{s+1}})$. As usual, let $\mathrm{sl}_2(\mathbb{F}_p)$ be the set of all traceless 2×2 matrices with entries in $\mathbb{F}_p$.

In the sequel we shall use, without special reference, the formula

$$\det(A - xI) = x^2 - x\mathrm{Tr}(A) + \det(A).$$

We are going now to define a map $\varphi_s \colon \mathrm{sl}_2(\mathbb{F}_p) \to \mathrm{SL}_2(\mathbb{Z}/(p^{s+1}))$ that will play a crucial role in what follows. Given $M \in \mathrm{sl}_2(\mathbb{F}_p)$, we pick a matrix $\widetilde{M}$ with entries in $\mathbb{Z}$ whose reduction modulo p is M, and we define $\varphi_s(M) = I + p^s\widetilde{M}$. Clearly, $\varphi_s(M)$ is independent of the particular choice of $\widetilde{M}$; slightly abusing the notation, we shall often write $1 + p^s M$ instead of $I + p^s\widetilde{M}$.

The following property is obvious.

Proposition 7.2.1 *We have a short exact sequence*

$$\mathrm{sl}_2(\mathbb{F}_p) \xhookrightarrow{\varphi_s} \mathrm{SL}_2(\mathbb{Z}/(p^{s+1})) \xrightarrow{\pi_s} \mathrm{SL}_2(\mathbb{Z}/(p^s)). \tag{7.2}$$

Note that, by restriction to the subgroup G_{s+1} of $\mathrm{SL}_2(\mathbb{Z}/(p^{s+1}))$ we obtain a short exact sequence

$$V_s \overset{\varphi_s}{\hookrightarrow} G_{s+1} \overset{\pi_s}{\twoheadrightarrow} G_s, \tag{7.3}$$

where $V_s = \varphi_s^{-1}(K_{s+1})$ is a subspace of $\mathrm{sl}_2(\mathbb{F}_p)$.

Consider the chain of projections $G_s \twoheadrightarrow G_{s-1} \twoheadrightarrow \ldots \twoheadrightarrow G_2 \twoheadrightarrow G_1$. It produces a sequence $V_{s-1}, \ldots, V_2, V_1$ of subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$. We shall now see some relations among them.

Proposition 7.2.2 *If $p^s \neq 2$ then $V_s \subset V_{s+1}$. If $p^s = 2$ and $-I \in \Gamma$ then $I \in V_1$ and $V_1 \subset V_2 + (I)$.*

Proof – Let M be an element of V_s, so that G_{s+1} contains the element $I + p^s M$. By surjectivity of the projection $\pi_{s+1}\colon G_{s+2} \to G_{s+1}$, there exists a matrix N with entries in $\mathbb{Z}/(p^2)$ such that $I + p^s N \in G_{s+2}$ projects to $I + p^s M$; obviously, $N \equiv M \pmod{p}$. In G_{s+2} the p-th power of $I + p^s N$ is

$$(I+p^sN)^p = \binom{p}{0} I + \binom{p}{1} p^s N + \binom{p}{2} p^{2s} N^2 = I + p^{s+1}\left(M + \binom{p}{2} p^{s-1} M^2\right),$$

implying that $M + \binom{p}{2} p^{s-1} M^2$ lies in V_{s+1}.

If $p \neq 2$ or $s > 1$, then p divides $\binom{p}{2} p^{s-1}$ and therefore $M \in V_{s+1}$.

If $p = 2$ and $s = 1$ then $M + M^2$, that is $M + I\det(M)$, lies in V_2. Note that, for $p = 2$, the assumption $-I \in \Gamma$ implies $I + 2I = -I \in G_2$, so that $I \in V_1$. □

Corollary 7.2.3 *Let Γ be a congruence subgroup of the exact level p^e. If $V_s = \mathrm{sl}_2(\mathbb{F}_p)$ for some s, then $e \leqslant s$.*

Proof – It suffices to show that $V_s = \mathrm{sl}_2(\mathbb{F}_p)$ implies $V_{s+1} = \mathrm{sl}_2(\mathbb{F}_p)$. This follows from the above proposition if $p^s > 2$, and it is verified by inspection for $p^s = 2$. □

Let us denote by γ_A the conjugation $\gamma_A(M) = A^{-1}MA$ by an invertible matrix A. Since the trace of a matrix is invariant under conjugation, γ defines an action of G_1 on $\mathrm{sl}_2(\mathbb{F}_p)$.

Proposition 7.2.4 *The space V_s is G_1-stable.*

Proof – Since $\varphi(V_s) = K_{s+1}$ is a normal subgroup of G_{s+1}, the space V_s is G_{s+1}-stable. We conclude by surjectivity of the projection $G_{s+1} \to G_1$. □

Summarizing the above results, the kernels K_{s+1} of the subsequent projections $G_{s+1} \to G_s$ correspond, with the possible exception of K_2 when $p = 2$, to a nested chain of G_1-stable subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$. Note that $\mathrm{sl}_2(\mathbb{F}_p)$ has dimension 3 over $\mathbb{F}_p$; thus any nested chain of subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$ can contain no more than two distinct non-trivial proper elements.

Under the requirements of (7.1), we can give a restriction on V_1.

Proposition 7.2.5 *If Γ satisfies (7.1), then $\#G_2 \geqslant (p^4 - p^2)/2$ and $V_1 \neq \langle O \rangle$. If moreover $[\mathrm{SL}_2(\mathbb{F}_p) : G_1] > 2$, then $\dim(V_1) \geqslant 2$.*

Proof – Let μ be the index of G_1 in $\mathrm{SL}_2(\mathbb{F}_p)$. Then G_1 has cardinality

$$\#G_1 = \#\mathrm{SL}_2(\mathbb{F}_p)/\mu = (p^3 - p)/\mu.$$

Under the condition (7.1), Propositions 5.1.5 and 5.1.6 imply

$$\#G_2 \geqslant \#M_{p^2}/2 = (p^4 - p^2)/2;$$

then $V_1 \cong K_2 \cong G_2/G_1$ has cardinality

$$\#V_1 \geqslant \frac{(p^4 - p^2)/2}{(p^3 - p)/\mu} = p\mu/2.$$

For $p > 2$ we have $p\mu/2 > 1$, while for $p = 2$ we have $V_1 \ni I$ by Proposition 7.2.2; in both cases $V_1 \neq \langle O \rangle$.

If moreover $\mu > 2$ then $p\mu/2 > p$ and $\dim(V_1) > 1$. □

We conclude this subsection with yet another relation between the spaces V_s.

Proposition 7.2.6 *Let M_1 and M_2 be elements of V_s. Then $M_1M_2 - M_2M_1$ lies in V_{2s}.*

Proof – For $i = 1, 2$ we fix an integer matrix N_i, defined modulo p^{s+1}, such that $X_i = I + p^s N_i \in G_{2s+1}$ projects to $I + p^s M_i \in G_{s+1}$; obviously, $N_i \equiv M_i \pmod{p}$. Note that

$$X_1X_2 - X_2X_1 \equiv p^{2s}(N_1N_2 - N_2N_1) \qquad (\mathrm{mod}\ p^{2s+1}).$$

Then the commutator of X_1 and X_2 is

$$X_1X_2(X_2X_1)^{-1} = I + p^{2s}(N_1N_2 - N_2N_1),$$

which concludes the proof. □

7.2.2 The lifting

We want to determine all possible groups G_{s+1} projecting on some fixed G_s. In this subsection we describe our strategy in general terms. In the subsequent sections we apply it in concrete situations.

As we have seen, the possible kernels K_{s+1} are described by special subspaces V_s of $\mathrm{sl}_2(\mathbb{F}_p)$; fix one of them. We also fix a set $\{X\}$ of generators of the group G_s.

By choosing a lifting $\tilde{X}$ in $\mathrm{SL}_2(\mathbb{Z}/(p^{s+1}))$ for every generator X from the fixed set of generators, and by taking the smallest group that contains all of these liftings and K_{s+1} we obtain a candidate $\tilde{G}_s$ for G_{s+1}. This group will project on G_s, but the kernel of the projection can possibly be larger than K_{s+1}.

Now let w be a word over the fixed set of generators $\{X\}$. It represents an element R_w in G_s; it also represents, replacing every X with the corresponding lifting $\tilde{X}$, an element $\tilde{R}_w$ in[1] $\tilde{G}_s = \langle \{\tilde{X}\}, K_{s+1} \rangle$, which projects onto R_w. The

[1] Notice that $\tilde{R}_w$ depends on the word w, not only on the element R_w; we may well have $R_{w_1} = R_{w_2}$ and $\tilde{R}_{w_1} \neq \tilde{R}_{w_2}$.

kernel of the projection $\tilde{G}_s \to G_s$ is K_{s+1} if and only if $\tilde{R}_w \in K_{s+1}$ for every w such that $R_w = I$.

For every X in G_s we fix, once and for all, a lifting $\tilde{X}'$. Then any other lifting $\tilde{X}$ of X (a "variable" lifting) is of the form $\tilde{X} = \tilde{X}'(I + p^s M_X)$ for some[2] M_X in $\mathrm{sl}_2(\mathbb{F}_p)$. Any word w now represents an element $R_w \in G_s$ and two liftings of R_w, namely $\tilde{R}'_w \in \langle\{\tilde{X}'\}\rangle$ and $\tilde{R}_w \in \langle\{\tilde{X}\}\rangle$; note that $\tilde{R}_w$ is of the form $\tilde{R}_w = \tilde{R}'_w(I + p^s M_w)$, for some[3] $M_w \in \mathrm{sl}_2(\mathbb{F}_p)$. Denoting by γ_A the conjugation map $M \mapsto A^{-1}MA$ for A in G_s, we can explicitly write every M_w in terms of the M_X (and of the X) by recursion on the length of w, via the cocycle relation

$$M_{w_1 w_2} = \gamma_{R_{w_2}}(M_{w_1}) + M_{w_2}.$$

For every word w such that $R_w = I$ we compute the corresponding $\tilde{R}'_w$, which lies in[4] $\varphi_s(\mathrm{sl}_2(\mathbb{F}_p))$; then the condition $\tilde{R}_w \in K_{s+1} = \varphi_s(V_s)$ can be written as

$$\varphi_s^{-1}(\tilde{R}'_w) + M_w \in V_s. \tag{7.4}$$

Note that, since $K_{s+1} = \varphi_s(V_s)$, the group $\tilde{G}_s = \langle\{\tilde{X}\}, K_{s+1}\rangle$ is determined by the cosets $M_X + V_s$ and (7.4) becomes an equation over $\mathrm{sl}_2(\mathbb{F}_p)/V_s$ in the unknowns $M_X + V_s$.

Now let W be a set of words such that the group G_s is defined by the fixed set of generators $\{X\}$ and the relations from W. Then it suffices to verify condition (7.4) only for the words $w \in W$. In the concrete examples below, we shall choose the sets of generators and relations in the most convenient way.

With as abuse of notation, we shall often denote by the same letter both a generator X of G_s and its fixed lifting $\tilde{X}'$.

We are now ready to begin our inspection of groups of prime power level. From time to time, the requirement (7.1) will provide more restrictions. We shall begin with the groups such that $p \neq 2$ divides the order of G_1, then turn to those such that $p \neq 2$ does not divide the order of G_1, and finally consider the case $p = 2$.

We fix for $\mathrm{sl}_2(\mathbb{F}_p)$ the basis

$$A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \quad B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \quad C = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right). \tag{7.5}$$

7.3 The triangular case

Throughout this and the following section we shall assume $p \neq 2$. In this section we consider groups G_s that satisfy (7.1) and such that p divides the order of G_1.

As we have seen in the previous chapter, G_1 is either $\mathrm{SL}_2(\mathbb{F}_p)$ or conjugate to $\mathrm{ST}_2(\mathbb{F}_p)$. In the latter case we can assume $G_1 = \mathrm{ST}_2(\mathbb{F}_p)$ and consider only the cases $p = 3, 5, 7, 13$, due to the requirements of (7.1).

We begin by studying the adjoint representations of $\mathrm{ST}_2(\mathbb{F}_p)$ and $\mathrm{SL}_2(\mathbb{F}_p)$, in order to find the subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$ that are stable under their action.

Fix a generator g of the multiplicative group $\mathbb{F}_p^*$ and consider the matrices $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, and $X = \left(\begin{smallmatrix} g^{-1} & 0 \\ 0 & g \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{F}_p)$. The element T generates

[2] The notation M_X is slightly abusive; in fact M_X depends on the variable lifting $\tilde{X}$, so it would be more correct to write $M_{\tilde{X}}$.

[3] Again M_w depends not only on w but also on the variable lifting.

[4] See Subsection 7.2.1 for the definition of φ_s.

the maximal unipotent group $\{\left(\begin{smallmatrix}1 & * \\ & 1\end{smallmatrix}\right)\}$; the elements T and X, together, generate the special triangular group $\mathrm{ST}_2(\mathbb{F}_p)$; the three elements S, T, and X generate the special linear group[5] $\mathrm{SL}_2(\mathbb{F}_p)$.

We use the notation (7.5). Recall that in this section we assume that $p > 2$.

Proposition 7.3.1 *The only non-trivial* $\mathrm{ST}_2(\mathbb{F}_p)$*-invariant subspaces of* $\mathrm{sl}_2(\mathbb{F}_p)$ *are* $\langle B\rangle$ *and* $\langle A, B\rangle$*. There are no non-trivial,* $\mathrm{SL}_2(\mathbb{F}_p)$*-invariant subspaces of* $\mathrm{sl}_2(\mathbb{F}_p)$*.*

Proof – We consider the basis

$$e_1 = 4B, \quad e_2 = 2A, \quad e_3 = -A - 2C$$

for $\mathrm{sl}_2(\mathbb{F}_p)$. In this basis, the conjugation map $M \mapsto T^{-1}MT$ has the matrix

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

Hence the non-trivial T-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$ are $\langle e_1\rangle = \langle B\rangle$ and $\langle e_1, e_2\rangle = \langle A, B\rangle$. Since both are also X-invariant, they are $\mathrm{ST}_2(\mathbb{F}_p)$-invariant, and there are no other. Since none of them is S-invariant, the group $\mathrm{SL}_2(\mathbb{F}_p)$ has no non-trivial invariant subspaces in $\mathrm{sl}_2(\mathbb{F}_p)$. □

This proposition allows us to settle the case $G_1 = \mathrm{SL}_2(\mathbb{F}_p)$ (when $p > 2$).

Corollary 7.3.2 *If* Γ *satisfies* (7.1) *and* $G_1 = \mathrm{SL}_2(\mathbb{F}_p)$*, then* $\Gamma = \mathrm{SL}_2(\mathbb{Z})$*.*

Proof – Straightforward from Proposition 7.2.5. □

We consider now a group $G_2 < \mathrm{SL}_2(\mathbb{Z}/(p^2))$ that satisfies (7.1) and that projects on $G_1 = \mathrm{ST}_2(\mathbb{F}_p)$, for some prime $p > 2$. The index of $\mathrm{ST}_2(\mathbb{F}_p)$ in $\mathrm{SL}_2(\mathbb{F}_p)$ is p; using Proposition 7.2.5 we obtain that V_1 has dimension at least 2, i.e. that V_1 is either $\langle A, B\rangle$ or $\mathrm{sl}_2(\mathbb{F}_p)$. In the second case, Γ would have exact level p, which is treated in the previous chapter. Hence $V_1 = \langle A, B\rangle$.

The group $G_1 = \mathrm{ST}_2(\mathbb{F}_p)$ is generated by T and X, with the relations

$$R_1 = T^p = I, \qquad R_2 = X^{p-1} = I, \qquad R_3 = X^{-1}TXT^{-g^2} = I,$$

where g is the generator of $\mathbb{F}_p$ fixed in the beginning of this section[6]. Fix a lifting of g in $\mathbb{Z}/(p^2)^*$; by abuse of notation, we denote it as g as well. We also fix in $\mathrm{SL}_2(\mathbb{Z}/(p^2))$ the liftings $T = \left(\begin{smallmatrix}1 & 1 \\ 0 & 1\end{smallmatrix}\right)$ and $X = \left(\begin{smallmatrix}g^{-1} & 0 \\ 0 & g\end{smallmatrix}\right)$. The group G_2 is generated by $I + pA$, $I + pB$, $\tilde{T} = T(I + ptC)$, and $\tilde{X} = X(I + pxC)$, for some t and x in $\mathbb{F}_p$.

We shall prove the following proposition.

Proposition 7.3.3 *Let* $p \neq 2$ *be a prime and let* Γ *be a congruence subgroup of exact level* p^s*, with* $s > 1$*, which satisfies* (7.1) *and with* G_1*,* K_2*, and* G_2

[5] Actually, already S and T generate $\mathrm{SL}_2(\mathbb{F}_p)$, but it is more convenient for us to include X in the set of generators.

[6] Since $T^p = 1$, the matrix T^{-g^2} is well-defined.

as above. Then p is either 3 or 5. For $p = 3$ we have $x = 0$ and $t \neq 0$, and the groups G_2 defined by $t = 1$ and $t = -1$ are non-conjugate in $\mathrm{SL}_2(\mathbb{Z}/(9))$. For $p = 5$ we have $t \neq 0$, and the groups G_2 defined by x and by $t \neq 0$ lie in 4 distinct conjugacy classes, depending only on the choice of t. Moreover, any proper subgroup of G_2 has at least 3 orbits in $\mathcal{M}_{p^2}$.

Proof – Over $\mathrm{sl}_2(\mathbb{F}_p)/\langle A, B\rangle \cong \langle C\rangle$ the action of T is trivial and the action of X corresponds to multiplication by g^{-2}. Note that conjugation by $(I + pC)$ maps $\tilde{T}$ into $\tilde{T}K_2$ and maps $\tilde{X}$ to $\tilde{X}(I + p(1 - g^{-2})C)$; up to conjugation we can thus assume $x = 0$ when $g^2 \not\equiv 1 \pmod{p}$, i.e. when $p \neq 3$.

We now compute the liftings in G_2 of the relations of G_1. All three $R_1 = T^p$, $R_2 = X^{p-1}$, and $R_3 = X^{-1}TXT^{-g^2}$ lie in K_2; we need to verify that M_{R_1}, M_{R_2}, and M_{R_3} lie in $\langle A, B\rangle$. Modulo $\langle A, B\rangle$ we have

$$\begin{aligned} M_{R_1} &= (t + \ldots + t)C = ptC = O; \\ M_{R_2} &= (1 + g^{-2} + \ldots + g^{-2(p-2)})xC; \\ M_{R_3} &= (g^{-2} - g^2)tC. \end{aligned}$$

The condition $M_{R_1} \in \langle A, B\rangle$ is always true. The condition $M_{R_2} = O$ is satisfied if $x = 0$, as is the case when $p \neq 3$, and implies $x = 0$ when $p = 3$; in any case we obtain $\tilde{X} = X$. The condition $M_{R_3} = O$ implies $t = 0$, unless the order of g in $\mathbb{Z}/(p)^*$ divides 4, i.e. unless p is either 3 or 5.

The group $G_1 = \mathrm{ST}_2(\mathbb{F}_p)$ has two orbits in M_p, namely the orbits of $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and of $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$; their stabilizers are $\langle T\rangle$ and $\{I\}$ respectively. This implies that the group G_2 has at least two orbits; there are precisely two of them if and only if the pull-back in G_2 of the stabilizer of an element of M_p acts transitively on the pull-back in M_{p^2} of that element. We have

$$\big(I + p(aA + bB)\big)\begin{pmatrix}1\\pz\end{pmatrix} = \begin{pmatrix}1+pa\\pz\end{pmatrix}, \quad \big(I + p(aA + bB)\big)\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}pb\\1+pa\end{pmatrix},$$

implying that the kernel K_2 acts transitively on the pull-back of $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, but has p distinct orbits in the pull-back of $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, corresponding to the elements $\left(\begin{smallmatrix}1\\pz\end{smallmatrix}\right)$, with $z \in \mathbb{F}_p$; the action of $\tilde{T}$ on these elements is

$$\tilde{T}\begin{pmatrix}1\\pz\end{pmatrix} = \begin{pmatrix}1+pz+pt\\pz+pt\end{pmatrix}.$$

This implies that G_2 has two orbits on M_{p^2} if and only if $t \neq 0$. As we have seen, this implies that p is either 3 or 5.

We conclude by inspection. □

We are now left with the cases $p = 3, 5$, where the values $t \neq 0$ correspond to the distinct conjugacy classes of G_2. When $p = 3$ we have $x = 0$, while when $p = 5$ we can choose any x.

Proposition 7.3.4 *Let Γ be a congruence subgroup of exact level 5^s, for some $s > 1$, that projects modulo 5 on $\mathrm{ST}_2(\mathbb{F}_5)$. Then Siegel's theorem is effective for the couple (X_Γ, j).*

Proof – By Proposition 7.3.3 and Theorem 7.1.2 it will suffice to show that G_2 has a proper subgroup containing all elements of order at most 4.

The group $G_1 = \mathrm{ST}_2(\mathbb{F}_5)$ is generated by $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $X = \left(\begin{smallmatrix} -2 & 0 \\ 0 & 2 \end{smallmatrix}\right)$, with the relations $T^5 = I$, $X^2 = -I$, and $(XT)^2 = -I$. It contains no elements of trace -1 and its traceless elements are of the form $\pm XT^a$, for any $a \in \mathbb{F}_5$.

We fix $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $X = \left(\begin{smallmatrix} -7 & 0 \\ 0 & 7 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z}/(25))$. The group G_2 is generated by $M = I + 5A$, $N = I + 5B$, $\tilde{T} = T(I + 5tC)$, and $\tilde{X} = X(I + 5xC)$, where we can choose $x = t$. The relation $\tilde{T}^5 = N$ shows that the generator N is redundant.

In G_2 there can be no elements of trace -1, since there are none in G_1, while the traceless elements of G_2 must be of the form $\pm\tilde{X}\tilde{T}^a M^b$, for some $a \in \mathbb{Z}/(25)$ and $b \in \mathbb{F}_5$. By means of direct computation we obtain that an element of the form $\tilde{X}\tilde{T}^a M^b = \tilde{X}\tilde{T}^a + 5bXT^a A$ has trace $15b$. This implies that the set of elements of G_2 od order at most 4 is $\{\pm\tilde{X}\tilde{T}^a\}$; by the relations $\tilde{T}^{25} = I$, $\tilde{X}^2 = -I$, and $(\tilde{X}\tilde{T})^2 = -I$, this set generates a proper subgroup of G_2, as required. □

Proposition 7.3.5 *Let Γ be a congruence subgroup of exact level 3^s, for some $s > 1$, that projects modulo 3 on $\mathrm{ST}_2(\mathbb{F}_3)$. Assume that Γ satisfies (7.1) and that there exists no congruence subgroup $\Gamma' \subsetneq \Gamma$ of exact level 3^s that contains all elliptic elements of Γ. Then $s \leqslant 3$ and the conjugacy class of Γ is uniquely determined by s. For $s = 3$ the curve X_Γ is of genus 1 (which implies Siegel's effectiveness by Theorem 4.2.2). For $s = 2$ the corresponding (X_Γ, j) is non-Siegelian.*

Proof – We fix $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z}/(9))$. By Proposition 7.3.3 the group G_2 is generated by $M = I + 3A$, $N = I + 3B$, and $-\tilde{T} = -T(I + 3tC)$, where t is either 1 or -1.

In G_2 there is no traceless element, since there is none in $G_1 = \mathrm{ST}_2(\mathbb{F}_3)$; the elements of G_2 with trace -1 must be of the form $(\tilde{T}M^a N^b)^{\pm 1}$. By means of direct computation we obtain

$$\mathrm{Tr}(\tilde{T}M^a N^b) = \mathrm{Tr}\ T + 3\mathrm{Tr}\ T \begin{pmatrix} a & b \\ t & -a \end{pmatrix} = 2 + 3(a + t - a) = -1 + 3(t + 1),$$

which implies that in G_2 there are neither traceless elements nor elements with trace -1, unless $t = -1$. Since no congruence subgroup $\Gamma' \subsetneq \Gamma$ contains all elliptic elements of Γ, we have $t = -1$ and $\tilde{T} = \left(\begin{smallmatrix} -2 & 1 \\ -3 & 1 \end{smallmatrix}\right)$. Note that $\tilde{T}^{-1}N\tilde{T} = N$ and $\tilde{T}^{-1}M\tilde{T} = MN^{-1}$, and that G_2 is stable under conjugation by $I + 3C = \left(\begin{smallmatrix} 1 & 0 \\ 3 & 1 \end{smallmatrix}\right)$.

We now consider G_3. Since $V_2 = V_1 = \langle A, B\rangle$, the kernel K_3 of $\pi_2 \colon G_3 \to G_2$ is generated by $I + 9A$ and $I + 9B$. We fix $T = \left(\begin{smallmatrix} -2 & 1 \\ -3 & 1 \end{smallmatrix}\right)$, $M = \left(\begin{smallmatrix} 13 & 0 \\ 0 & -2 \end{smallmatrix}\right)$, and $N = \left(\begin{smallmatrix} 1 & 3 \\ 9 & 1 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z}/(27))$. The group G_3 is generated by $I + 9A$, $I + 9B$ and, for some m, n, and t in $\mathbb{F}_p$, by the elements

$$\tilde{M} = M(I + 9mC), \qquad \tilde{N} = N(I + 9nC), \qquad -\tilde{T} = -T(I + 9tC).$$

We have the relations $\tilde{M}^3 = I + 9A$ and $\tilde{N}^3 = I + 9B$. Up to conjugation by $\left(\begin{smallmatrix} 1 & 0 \\ 3 & 1 \end{smallmatrix}\right)$ we may assume $m = 0$, and by a direct computation we obtain that $\tilde{N}\tilde{M}^{-1}\tilde{T}^{-1}\tilde{M}\tilde{T} \in K_3$ implies $n = 0$, i.e. $\tilde{N} = N$. Further computations on the

relations and on the elements with trace -1 show that $G_3 = \langle M, N, -T\rangle$ is unique up to conjugacy. We also have the relation $NM = MN^4$.

We now fix $M \equiv \left(\begin{smallmatrix} 40 & 0 \\ 0 & -2 \end{smallmatrix}\right)$ and $N \equiv \left(\begin{smallmatrix} 1 & 3 \\ 9 & 28 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z}/(81))$. Consider the matrices $\tilde{M} = M(I + 27mC)$ and $\tilde{N} = N(I + 27nC)$, with m and n in $\mathbb{F}_p$. For any choice of m and n we have

$$\tilde{M}^{-1}\tilde{N}^{-1}\tilde{M}\tilde{N}^4 = I - 27C \notin \langle I + 27A, I + 27B\rangle,$$

which implies that any subgroup G_4 of $\mathrm{SL}_2(\mathbb{Z}/(81))$ that projects on G_3 contains the kernel of the projection $\pi_3 \colon \mathrm{SL}_2(\mathbb{Z}/(81)) \to \mathrm{SL}_2(\mathbb{Z}/(27))$.

We conclude by a direct computation on the invariants for $s = 2$ and for $s = 3$. □

We can summarize the results of this section as follows.

Proposition 7.3.6 *Let Γ be a congruence subgroup of exact level p^s with $s > 1$ and $p \neq 2$, containing $-I$ and whose projection G_1 has order divisible by p (i.e. Γ is unipotent). Then Siegel's theorem is effective for (X_Γ, j), with the exception of one (up to conjugacy) subgroup of level 9, for which (X_Γ, j) is non-Siegelian.*

Proof – By Theorem 4.2.3 Siegel's theorem is effective for (X_Γ, j) if X_Γ has at least 3 cusps. As we have seen in the previous chapter this is the case unless $G_1 = \mathrm{SL}_2(\mathbb{F}_p)$ or $G_1 = \mathrm{ST}_2(\mathbb{F}_p)$, up to conjugacy. In the former case by Corollary 7.3.2 the curve X_Γ has at least 3 cusps and we conclude. In the latter case we have either $p = 3$ or $p = 5$ by Proposition 7.3.3, and we conclude by Propositions 7.3.4 and 7.3.5. □

7.4 The special cases

As in the previous section, we assume $p \neq 2$. In this section we consider groups G_s that satisfy (7.1) and such that p does not divide the order of G_1. As we have seen in the previous chapter, up to conjugacy there are ten possible groups G_1 for $p \neq 2$.

Note that the index of G_1 in $\mathrm{SL}_2(\mathbb{F}_p)$ is at least $p > 2$; by Proposition 7.2.5 this implies $\dim(V_1) \geqslant 2$.

We shall need a simple lemma, that will be used for $n = 3$, but we state the general case. It is certainly well-known, but we include a proof for the sake of completeness.

Lemma 7.4.1 *Let A be an algebra over a field of characteristic distinct from 2, and let $X_1, \ldots, X_n$ be invertible pairwise anti-commuting elements of A. Then $X_1, \ldots, X_n$ are linearly independent over the base field.*

Proof – Let $S = \sum_i a_i X_i$ be a linear combination of the X_i, with a_i in the base field. If $S = 0$ then for every i we have

$$0 = X_i S + S X_i = \sum_{j \neq i} a_i (X_i X_j + X_j X_i) + 2a_i X_i^2 = 2a_i X_i^2.$$

Since X_i is invertible in A and 2 is invertible in the base field, this implies that every coefficient a_i is 0. □

Now we have the following property, which allows us to immediately exclude seven of the ten cases referred to in the beginning of this section.

Proposition 7.4.2 *Let G be a subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$ and let $\bar{G}$ be its image in $\mathrm{PSL}_2(\mathbb{F}_p)$. If $\bar{G}$ contains a subgroup $\bar{H}$ isomorphic to $\mathbb{Z}/(2)\times\mathbb{Z}/(2)$, then a base of $\mathrm{sl}_2(\mathbb{F}_p)$ can be given by three elements of G.*

If $\bar{G}$ contains a subgroup isomorphic to the alternating group $\mathcal{A}_4$, then there are no non-trivial G-stable subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$.

Proof – Let $\bar{X}$ and $\bar{Y}$ be generators of $\bar{H}\equiv\mathbb{Z}/(2)\times\mathbb{Z}/(2)$ and let $\pm X$ and $\pm Y$ be their pullbacks in G. Since the elements X, Y, and XY are traceless, they belong to $\mathrm{sl}_2(\mathbb{F}_p)$. The relations

$$X^2=-I, \qquad Y^2=-I, \qquad (XY)^2=-I,$$

show that X, Y, and XY are pairwise anti-commutative; by the lemma above, they form a basis of $\mathrm{sl}_2(\mathbb{F}_p)$.

In this basis, the conjugation maps by X, Y, and XY have the matrices

$$\gamma_X=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix},\quad \gamma_Y=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&-1\end{pmatrix},\quad \gamma_{XY}=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix}.$$

This implies that the G-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$ are generated by subsets of $\{X,Y,XY\}$.

Let now $\bar{G}$ contain a subgroup isomorphic to A_4; in turn, this will contain a subgroup $\bar{H}$ isomorphic to the Klein group $T\cong\mathbb{Z}/(2)\times\mathbb{Z}/(2)$, and an element $\bar{R}$ that cycles, by conjugation, the non-trivial elements of $\bar{H}$. Taking a basis X, Y, and XY of $\mathrm{sl}_2(\mathbb{F}_p)$ as above, the pullback R of $\bar{R}$ in G cycles the spaces $\langle X\rangle$, $\langle Y\rangle$, and $\langle XY\rangle$. Thus the only G-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_p)$ are trivial. □

The above proposition enables us to exclude from our search the last four lines in the table of Section 6.2 (recall that $\mathcal{A}_4$ is both a subgroup of $\mathcal{A}_5$ and of $\mathcal{S}_4$). We are now left with three groups.

Proposition 7.4.3 *Let Γ be a congruence subgroup of exact level 5^s with $s>1$, that satisfies (7.1), and whose projection $\bar{G}_1$ in $\mathrm{PSL}_2(\mathbb{F}_5)$ is isomorphic to the dyhedral group $\mathcal{D}_3$. Then $s=2$ and Γ is unique up to conjugacy. The couple (X_Γ,j) is Siegelian, but Γ is generated by its elliptic elements and X_Γ has 2 cusps.*

Proof – Applying a suitable conjugation, we may assume that G_1 contains the elements $S=\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$, $R=\left(\begin{smallmatrix}0&2\\2&-1\end{smallmatrix}\right)$, and $SR=\left(\begin{smallmatrix}2&-1\\0&-2\end{smallmatrix}\right)$, which satisfy the relations

$$S^2=-I, \qquad R^3=I, \qquad (SR)^2=-I.$$

We also put $X=S^2R-SRS=\left(\begin{smallmatrix}-1&1\\1&1\end{smallmatrix}\right)$.

It is verified by inspection that the only non-trivial G_1-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_5)$ are $\langle S, SR\rangle$ and $\langle X\rangle$. By Proposition 7.2.5 we have $\dim(V_1) \geqslant 2$, thus $V_1 = \langle S, SR\rangle$. Further, since $X = S(SR) - (SR)S$, we have $X \in V_2$ by Proposition 7.2.6. This implies $V_2 = \mathrm{sl}_2(\mathbb{F}_p)$ and $s \leqslant 2$.

We fix $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $R = \left(\begin{smallmatrix} 0 & 7 \\ 7 & -1 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z}/(25))$. They satisfy

$$S^2 = -I, \qquad R^3 = I, \qquad (SR)^2 = -I.$$

The group G_2 is generated by $\tilde{S} = S(I + 5tX)$, $\tilde{R} = R(I + 5rX)$, $M = I + 5S$, and $N = I + 5SR$, for some t and r in $\mathbb{F}_5$. Up to conjugacy by $I + 5X$ we may assume $t = 0$, and since

$$\tilde{R}^3 = I + 15rX,$$

we also have $r = 0$. Thus $\tilde{S} = S$ and $\tilde{R} = R$.

It is verified by inspection that Γ is generated by its elliptic elements and that X_Γ is of genus 2 and has 2 cusps. □

Proposition 7.4.4 *Let Γ be a congruence subgroup of exact level 3^s with $s > 1$, that satisfies* (7.1) *and whose projection $G_1 < \mathrm{SL}_2(\mathbb{F}_3)$ has cardinality not divisible by* 3. *Then $s = 2$ and Γ contains no 3-elliptic elements. If there exists no congruence subgroup $\Gamma' \subsetneq \Gamma$ that contains all elliptic elements of Γ, then Γ is uniquely defined by G_1 up to conjugacy and the couple (X_Γ, j) is non-Siegelian.*

Proof – As we have seen in the preceeding chapter, $\bar{G}_1$ is isomorphic either to the cyclic group $\mathcal{C}_2 \cong \mathbb{Z}/(2)$ or to the dyhedral group $\mathcal{D}_2 \cong \mathbb{Z}/(2) \times \mathbb{Z}/(2)$. We take $R = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and $T = RS = \left(\begin{smallmatrix} -1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{F}_3)$; they satisfy

$$R^2 = -I, \qquad S^2 = -I, \qquad T^2 = -I.$$

Up to conjugacy we can assume that either $G_1 = \langle S\rangle$ or $G_1 = \langle R, S\rangle$.

By Proposition 7.4.2, the elements R, S, and T are a basis of $\mathrm{sl}_2(\mathbb{F}_3)$. By inspection, the 2-dimensional S-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_3)$ are:

$$\langle S, R\rangle, \qquad \langle S, T\rangle, \qquad \langle R, T\rangle, \qquad \langle S, R + T\rangle, \qquad \langle S, R - T\rangle;$$

among these only $\langle S, R\rangle$, $\langle S, T\rangle$, and $\langle R, T\rangle$ are also R-invariant.

Since Γ satisfies (7.1) the group G_2 has at most 2 orbit in M_9; if $G_1 = \langle S\rangle$ this implies $\#M_9 = 2\#G_2$, thus the stabilizer of any element in M_9 is trivial. The equations

$$\begin{aligned}(I + 3(S + R + T))\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} 1 & 0 \\ 3 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right),\\ (I + 3(S + R - T))\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) &= \left(\begin{smallmatrix} -2 & 3 \\ -3 & 4 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right),\end{aligned}$$

show that V_1 cannot be neither $\langle S, R + T\rangle$ nor $\langle S, R - T\rangle$.

Thus V_1 is one of $\langle S, R\rangle$, $\langle S, T\rangle$, or $\langle R, T\rangle$, for any G_1. Up to conjugacy we can assume $\langle R, S\rangle \subset V_1$. Since $T = SR - RS$, by Proposition 7.2.6 we have $T \in V_2$. Then $V_2 = \mathrm{sl}_2(\mathbb{F}_3)$ and $s \leqslant 2$.

We conclude by inspection. □

We can summarize the results of this section as follows.

Proposition 7.4.5 *Let Γ be a congruence subgroup of exact level p^s with $s > 1$ and $p \neq 2$, containing $-I$ and whose projection G_1 modulo p has cardinality not divisible by p. Then either Siegel's theorem is effective on (X_Γ, j) or the couple (X_Γ, j) is non-Siegelian, with the exception of a group of level 25 which (up to conjugacy) projects onto*

$$G_2 = \left\langle \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 7 \\ 7 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 5 \\ -5 & 1 \end{pmatrix}, \begin{pmatrix} 11 & -5 \\ 0 & -9 \end{pmatrix} \right\rangle.$$

Proof – By the results of the previous chapter, Siegel's theorem is effective for (X_Γ, j) unless G_1 is one of ten groups, up to conjugacy.

Asume $p = 3$. By Proposition 7.4.4 if $s > 2$ then X_Γ has at least 3 cusps; if $s = 2$ and Γ' contains all elliptic elements of Γ (i.e. 2-elliptic elements), then either $X_{\Gamma'}$ has at least 3 cusps, or $\Gamma' = \Gamma$ and (X_Γ, j) is one of the two non-Siegelian curves described in Proposition 7.4.4. We conclude by Theorem 7.1.2.

If $p = 5$ and $\bar{G}_1 \cong \mathcal{D}_3$, we conclude by Proposition 7.4.3 and Theorem 4.2.3.

If G_1 is one of the remaining seven groups then it contains a subgroup isomorphic to the fourth alternating group $\mathcal{A}_4$. By Proposition 7.4.2 we have $V_1 = \langle O \rangle$ since $s > 1$; by Proposition 7.2.5 X_Γ has at least 3 cusps, and we conclude by Theorem 4.2.3. □

7.5 The case $p = 2$

In this section we assume $p = 2$. Then G_1 is a subgroup of $\mathrm{SL}_2(\mathbb{F}_2) \cong \mathcal{S}_3$. Some proposition in this section are proved by inspection.

We take $R = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{F}_2)$, with $T = RS = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Note that $\mathrm{SL}_2(\mathbb{F}_2)$ is generated by R and S, with the relations

$$R^3 = I, \qquad S^2 = I, \qquad T^2 = I.$$

The following results are proved by inspection.

Proposition 7.5.1 *The elements I, S, and T are a basis for $\mathrm{sl}_2(\mathbb{F}_2)$. The non-trivial S-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_2)$ are:*

$$\langle I \rangle, \quad \langle S \rangle, \quad \langle I + S \rangle, \qquad \langle I, S \rangle, \quad \langle T, S \rangle, \quad \langle I + T, S \rangle.$$

The non-trivial R-invariant subspaces of $\mathrm{sl}_2(\mathbb{F}_2)$ are $\langle I \rangle$ and $\langle S, T \rangle$. □

Proposition 7.5.2 *Let $M \in V_1$. Then $M + M^2 \in V_2$.* □

Up to conjugacy, we may assume that G_1 is one of $\langle R \rangle$, $\langle S \rangle$, and $\mathrm{SL}_2(\mathbb{F}_2)$.

Proposition 7.5.3 *Let Γ be a congruence subgroup of exact order 2^s with $s > 1$, whose projection modulo 2 is a group of order 2. Assume that Γ satisfies (7.1), as well as any congruence subgroup $\Gamma' \subset \Gamma$ which contains all elliptic elements of Γ. Then $s \leqslant 5$ and Γ is uniquely determined by s up to conjugacy. For $s = 5$ the curve X_Γ has genus 1 (which implies Siegel's effectiveness by Theorem 4.2.2). For $2 \leqslant s \leqslant 4$ the couple (X_Γ, j) is non-Siegelian.*

Proof – We can assume $G_1 = \langle S \rangle$. By Proposition 7.2.2 we have $I \in V_1$ and by Proposition 7.2.5 we have $\dim(V_1) \geqslant 2$, since $[\mathrm{SL}_2(\mathbb{F}_2) : G_1] = 3$. By Proposition 7.5.1 and Corollary 7.2.3 this implies $V_1 = \langle I, S \rangle$. Since $S \in V_1$ we have $S^2 + S = I + S \in V_2$ by Proposition 7.5.2. Propositions 5.1.5 and 5.1.6 imply $\#G_3 \geqslant \#\mathcal{M}_8/2 = 24$, and since $\#G_2 = 8$ we have $\dim(V_2) \geqslant 2$. By Proposition 7.5.1 this implies $V_2 = \langle I, S \rangle$.

We verify by inspection that G_s is generated, for $s = 3, 4, 5$, by $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $M = \left(\begin{smallmatrix} 1 & -2 \\ -2 & -5 \end{smallmatrix}\right)$, and $N = \left(\begin{smallmatrix} 5 & -8 \\ -8 & 13 \end{smallmatrix}\right)$. For $s = 6$ and for any choice of $m, n \in \mathbb{F}_2$ the matrices $\tilde{M} = M(I + 32mT)$ and $\tilde{N} = N(I + 32nT)$ in $\mathrm{SL}_2(\mathbb{Z}/(64))$ satisfy

$$\tilde{M}^{-9}\tilde{N}^{-1}\tilde{M}\tilde{N}^5 = I + 32(I + T),$$

which implies $I + T \in V_5$. Since $V_2 = \langle I, S \rangle \subset V_5$ by Proposition 7.2.2, we obtain $V_5 = \mathrm{sl}_2(\mathbb{F}_2)$ and $s \leqslant 5$.

We conclude by inspection on the genus and the number of cusps of X_Γ. □

Proposition 7.5.4 *Let Γ be a congruence subgroup of exact level 2^s with $s > 0$, whose projection modulo 2 is a group of order 3. Assume that Γ satisfies (7.1), as well as any congurence subgroup $\Gamma' \subset \Gamma$ which contains all elliptic elements of Γ. Then $s \leqslant 4$ and Γ is uniquely determined by s up to conjugacy. Moreover the couple (X_Γ, j) is non-Siegelian.*

Proof – We can assume $G_1 = \langle R \rangle$. By Proposition 7.2.2 we have $-I \in V_1$; by Proposition 7.5.1 and Corollary 7.2.3 this implies $V_1 = \langle I \rangle$. Propositions 5.1.5 and 5.1.6 imply $\#G_3 \geqslant \#\mathcal{M}_8/2 = 24$, and since $\#G_2 = 6$ we have $\dim(V_2) \geqslant 2$. By Proposition 7.5.1 this implies $V_2 = \langle S, T \rangle$. Finally, V_4 contains $ST - TS = I$ by Proposition 7.2.6; since $V_2 = \langle S, T \rangle \subset V_4$ by Proposition 7.2.2, this implies $V_4 = \mathrm{sl}_2(\mathbb{F}_2)$ and $s \leqslant 4$.

We conclude by inspection. □

Proposition 7.5.5 *Let Γ be a congruence subgroup of exact level 2^s, for some $s > 1$, that projects modulo 2 on $\mathrm{SL}_2(\mathbb{F}_2)$ and that satisfies (7.1). Then $s \leqslant 4$ and Γ belongs to one of eight distinct conjugacy classes. Moreover the couple (X_Γ, j) is non-Siegelian.*

Proof – We have $G_1 = \mathrm{SL}_2(\mathbb{F}_2) = \langle R, S \rangle$. By Proposition 7.2.5 we have $-I \in V_1$, which implies $V_1 = \langle I \rangle$ by Proposition 7.5.1 and Corollary 7.2.3. Propositions 5.1.5 and 5.1.6 imply $\#G_3 \geqslant \#\mathcal{M}_8/2 = 24$, and since $\#G_2 = 12$ we have $\dim(V_2) \geqslant 1$. By Proposition 7.5.1 this implies either $V_2 = \langle I \rangle$ or $V_2 = \langle S, T \rangle$.

If $V_2 = \langle I \rangle$ then $\#G_3 = 24$, and since Propositions 5.1.5 and 5.1.6 imply $\#G_4 \geqslant \#\mathcal{M}_{16}/2 = 96$ we obtain $\dim(V_3) \geqslant 2$. By Proposition 7.5.1 this implies $\langle S, T \rangle \subset V_3$, and since $\langle I \rangle = V_2 \subset V_3$ by Proposition 7.2.2 we obtain $V_3 = \mathrm{sl}_2(\mathbb{F}_p)$ and $s \leqslant 3$.

If $V_2 = \langle S, T \rangle$ then Proposition 7.2.6 implies $I = ST - TS \in V_4$, and since $V_2 \subset V_4$ by Proposition 7.2.2, we obtain $V_4 = \mathrm{sl}_2(\mathbb{F}_2)$ and $s \leqslant 4$. (In this case G_3 can belong to two distinct conjugacy classes in $\mathrm{SL}_2(\mathbb{Z}/(8))$ and G_4 can belong to four distinct conjugacy classes in $\mathrm{SL}_2(\mathbb{Z}/(16))$.)

We conclude by inspection. □

We summarize the results of this section.

Proposition 7.5.6 *Let Γ be a congruence subgroup of exact level 2^s with $s > 1$, containing $-I$. Then either Siegel's theorem is effective on (X_Γ, j) or the couple (X_Γ, j) is non-Siegelian.*

Proof – If Γ does not satisfy the hypotesis of any of the Propositions 7.5.3, 7.5.4, or 7.5.5, then there exists a congruence subgroup $\Gamma' \subset \Gamma$ that contains the elliptic elements of Γ and such that $X_{\Gamma'}$ has at least 3 cusps. We conclude by Theorem 7.1.2. □

Table 7.1: Non-Siegelian modular curves X_Γ of pure prime power level p^s

G_s	p^s	μ	ν_∞	ν_2	ν_3	$\mathbf{g}$	
$G_{4.1}$	4	6	2	2	0	0	
$G_{8.1}$	8	12	2	4	0	0	
$G_{16.1}$	16	24	2	8	0	0	
$G_{4.2}$	4	8	2	0	2	0	
$G_{8.2}$	8	16	2	0	4	0	
$G_{16.2}$	16	32	2	0	8	0	
$G_{4.3}$	4	4	1	2	1	0	
$G_{8.3}{}^\dagger$	8	8	1	2	2	0	2 groups
$G_{16.3}{}^\dagger$	16	16	1	2	4	0	4 groups
$G_{8.4}$	8	16	2	4	1	0	
$G_{9.1}$	9	12	2	0	3	0	
$G_{9.2}$	9	18	2	6	0	0	
$G_{9.3}$	9	9	1	5	0	0	

Chapter 8

The mixed level case

8.1 Introduction

In this chapter we study groups of mixed level. We shall prove the following result.

Theorem 8.1.1 *Let Γ be a congruence subgroup of level not dividing the number $2^{21}\cdot 3^7\cdot 5^3\cdot 7^2\cdot 11\cdot 13$. Then Siegel's theorem is effective for the pair (X_Γ, j).*

Let Γ be a congruence subgroup of exact level n, and let the factorization of n be $n = \prod_{i\in I} q_i = \prod_{i\in I} p_i^{e_i}$, where the p_i are distinct primes and $e_i > 0$ for every $i \in I$.

For every positive integer d we denote by Γ_d the composite group $\Gamma\cdot\Gamma(d)$, of level dividing d, and by $\mathcal{G}_d < \mathrm{SL}_2(\mathbb{Z}/(d))$ its projection modulo d. Note that if Siegel's theorem is effective for (X_{Γ_d}, j) then it is effective also for (X_Γ, j).

The group $\mathrm{SL}_2(\mathbb{Z}/(n))$ is isomorphic to the direct product $\prod_{i\in I}\mathrm{SL}_2(\mathbb{Z}/(q_i))$; this allows us to consider $\mathcal{G} = \mathcal{G}_n$ as a subgroup of the direct product $\prod_{i\in I}\mathcal{G}_{q_i}$.

Remark 2 *Probably, the assumption on the level in Theorem 8.1.1 can be relaxed, but at the moment, the methods of the present thesis do not allow treatment of all possible Siegelian modular curves of mixed level. Consider, for instance, two congruence subgroups Γ_5 and Γ_7 of exact levels 5 and 7, whose projections in $\mathrm{PSL}_2(\mathbb{F}_p)$ (see Table 6.1) are isomorphic to the fourth alternating group $\mathcal{A}_4$ and to the fourth symmetric group $\mathcal{S}_4$, respectively; their intersection $\Gamma_5\cap\Gamma_7$ is a congruence subgroup Γ of the exact level 35, generated by its elliptic elements and such that X_Γ has genus 2 and only one cusp. The couple (X_Γ, j) is non-Siegelian, but eludes our methods.*

8.2 Proof of Theorem 8.1.1

We begin with the following useful observation.

Proposition 8.2.1 *Let $\{S_i\}_{i\in I}$ be a finite family of finite groups S_i and let T be a subgroup of the (formal) direct product $S = \prod_{i\in I} S_i$. For every $J \subset I$ we*

have a natural projection $\pi_J\colon S \to S_J$ *and a natural embedding* $\iota_J\colon S_J \to S$, *where* $S_J = \prod_{i\in J} S_i$. *Let* T_J *and* U_J *be the subgroups of* S_J *defined by*

$$T_J = \pi_J(T), \qquad\qquad \iota_J(U_J) = T \cap \iota_J(S_J).$$

Then U_J *is a normal subgroup of* T_J. *Let also* r_i *be the index of* $U_{\{i\}}$ *in* $T_{\{i\}}$. *Then* r_j *divides* $\prod_{i\neq j} r_i$ *for every* $j \in I$.

Proof – Let $\{J, K\}$ be a partition of I. The group $\iota_J(U_J) = \mathrm{Ker}(\pi_K|_T)$ is normal in T; then $U_J = \pi_J \circ \iota_J(U_J)$ is a normal subgroup of $T_J = \pi_J(T)$. The composite map $T \twoheadrightarrow T_J \twoheadrightarrow T_J/U_J$ has kernel $U_J \times U_K$ and induces an isomorphism $T/(U_J \times U_K) \cong T_J/U_J$, which proves $T_J/U_J \cong T_K/U_K$.

Now note that

$$\prod_{i\in K} U_{\{i\}} < U_K < T_K < \prod_{i\in K} T_{\{i\}}.$$

This implies that $\#(T_K/U_K)$ divides $\#(\prod_{i\in K} T_{\{i\}})/(\prod_{i\in K} U_{\{i\}}) = \prod_{i\in K} r_i$. Taking $J = \{j\}$ and $K = I - J$ we obtain $T_{\{j\}}/U_{\{j\}} \cong T_K/U_K$, hence the result. □

Applying the above proposition to the group $\mathcal{G} < \prod_{i\in I} \mathcal{G}_i$ we immediately obtain the following result.

Corollary 8.2.2 *Let* Γ *be a congruence subgroup of exact level* $n = \prod_{i\in I} q_i$. *Then for every* $i \in I$ *the congruence subgroup* $(\Gamma \cap \Gamma(n/q_i)) \cdot \Gamma(q_i)$ *of exact level* q_i *projects modulo* q_i *onto a normal subgroup* $\mathcal{H}_{q_i}$ *of* $\mathcal{G}_{q_i}$ *of index* r_i, *and* r_j *divides* $\prod_{i\neq j} r_i$ *for every* $j \in I$. □

The following result is certainly well-known, but we include a proof for the sake of completeness.

Proposition 8.2.3 *Let* p *be a prime and let* H_s *be a normal subgroup of* $\mathrm{SL}_2(\mathbb{Z}/(p^s))$ *for some integer* $s > 0$. *If* $H_s \neq \mathrm{SL}_2(\mathbb{Z}/(p^s))$ *then* p *divides the index of* H_s.

Proof – When $s = 1$, the normal subgroup $H = H_1$ of $\mathrm{SL}_2(\mathbb{F}_p)$ is union of conjugacy classes. It is easily verified that all elements with fixed trace $t \neq \pm 2$ lie in the same class. If p does not divide the index of H then p divides the cardinality of H, which contains a cyclic subgroup of order p conjugate to $\left\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\right\rangle$. Since $\left\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\right\rangle$ and $\left\langle \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\right\rangle$ are conjugate, it follows that for every $x, y \in \mathbb{F}_p$ the matrix $\left(\begin{smallmatrix} 1 & y \\ x & 1+xy \end{smallmatrix}\right)$ of trace $2 + xy$ lies in H, as all their conjugates. We conclude $H = \mathrm{SL}_2(\mathbb{F}_p)$.

For $s > 1$, we the projection of H_s modulo p^{s-1} is a normal subgroup H_{s-1} of $\mathrm{SL}_2(\mathbb{Z}/(p^{s-1}))$. We have

$$[\mathrm{SL}_2(\mathbb{Z}/(p^s)) : H_s] = p^a [\mathrm{SL}_2(\mathbb{Z}/(p^{s-1})) : H_{s-1}]$$

for some $a \geqslant 0$. We conclude by induction. □

We immediately deduce the following result.

Proposition 8.2.4 *Let* Γ *be congruence subgroup of exact level* n *and let* $p > 3$ *be the largest prime appearing in the factorization of* n. *Then* $\mathcal{G}_p \neq \mathrm{SL}_2(\mathbb{F}_p)$.

Proof – Let $q = p^e$, $\mathcal{G}_q$, and $\mathcal{H}_q$ be as in Corollary 8.2.2. Since p does not divide $\#\mathrm{SL}_2(\mathbb{Z}/(p'^{e'})) = (p'+1)(p'-1)p'^{3e'-2}$ for any prime $p' < p$, it cannot divide $[\mathcal{G}_q : \mathcal{H}_q]$ by Corollary 8.2.2. By the above proposition, this implies that if $\mathcal{G}_p = \mathrm{SL}_2(\mathbb{F}_p)$ then $\mathcal{H}_q = \mathcal{G}_q$, but in this case p would not divide n. □

Corollary 8.2.5 *Let Γ be congruence subgroup of exact level n and let $p > 13$ be the largest prime appearing in the factorization of n. Then Siegel's theorem is effective for (X_Γ, j).*

Proof – Consider the congruence subgroup Γ_p of level dividing $p > 13$. As we have seen in the previous chapters, either Siegel's theorem is effective for (X_{Γ_p}, j) or $\mathcal{G}_p = \mathrm{SL}_2(\mathbb{F}_p)$. We exclude the latter case by the above proposition. □

The following result is obvious.

Proposition 8.2.6 *Let Γ_p be a congruence subgroup of exact level p^e and let Γ'_p be a congruence subgroup of exact level $p^{e'}$ with $\Gamma_p < \Gamma'_p$, where $e \geqslant e' \geqslant 0$ are integers and p is a prime. Then the index $[\Gamma'_p : \Gamma_p]$ divides $p^{3e'-2}(p+1)(p-1)$ and is divisible by $p^{e-e'}$.* □

Proof of Theorem 8.1.1 – Let Γ be a subgroup of exact level $n = \prod p^{e_p}$. If the set of primes p is not contained in $\{2, 3, 5, 7, 11, 13\}$ then we conclude by the above corollary. Assume now that n factors in the primes $2, 3, 5, 7, 11, 13$.

By Corollary 8.2.2 for every prime p the congruence subgroups $\Gamma'_p = \Gamma \cdot \Gamma(p^{e_p})$ and $\Gamma_p = (\Gamma \cap \Gamma(n/p^{e_p})) \cdot \Gamma(p^{e_p})$ of exact levels $p^{e'_p}$ and p^{e_p} respectively project modulo p^e on subgroups $\mathcal{G}_{p^{e_p}}$ and $\mathcal{H}_{p^{e_p}}$ of $\mathrm{SL}_2(\mathbb{Z}/(p^{e_p}))$, with $\mathcal{H}_{p^{e_p}} \lhd \mathcal{G}_{p^{e_p}}$.

If Siegel's theorem is effective for $(X_{\Gamma'_p}, j)$ then it is effective for (X_Γ, j), too. Otherwise, by the results of the previous chapter, we have $e'_2 \leqslant 4$, $e'_3, e'_5 \leqslant 2$, and $e'_7, e'_{11}, e'_{13} \leqslant 1$. Applying the above proposition together with Corollary 8.2.2 we obtain that n divides $2^{21} \cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 11 \cdot 13$. □

Bibliography

[1] P. Alvanos and Yu. Bilu and D. Poulakis, Characterizing algebraic curves with infinitely many integral points, *Int. J. Number Theory*, (...) 2008.

[2] A. Baker and J. Coates, Integer points on curves of genus 1, *Proc. Cambridge Philos. Soc.* **67** (1970), 595–602.

[3] Yu. F. Belotserkovskiĭ (Yu. Bilu), Effective analysis of a new class of Diophantine equations. (Russian. English summary) *Vestsī Akad. Navuk BSSR Ser. Fī z.-Mat. Navuk* **125**, (1988), no. 6 34–39.

[4] Yu. Bilu, Effective analysis of integral points on algebraic curves, Ph. D. thesis, Beer-Sheva, 1993.

[5] Yu. Bilu, Effective analysis of integral points on algebraic curves, *Israel J. Math.* **90** (1995), 235–252.

[6] Yu. Bilu, Baker's method and modular curves, *A Panorama of Number Theory or The View from Baker's Garden* (edited by G. Wustholz), 73–88, Cambridge University Press, 2002.

[7] J. W. S. Cassels, *An introduction to the Geometry of Numbers*, Springer, 1997.

[8] R. Dvornicich and U. Zannier, Fields containing values of algebraic functions. II. (On a conjecture of Schinzel). *Acta Arith.* **72** (1995), no. 3, 201–210.

[9] R. Dvornicich and U. Zannier, Local-global divisibility of rational points in some commutative algebraic groups, *Bull. Soc. Math. France* **129** (2001), no. 3, 317–338.

[10] R. Dvornicich and U. Zannier, On a local-global principle for the divisibility of a rational point by a positive integer, *Bull. London Math. Soc.* **39** (2007), 27–34.

[11] D. S. Kubert and S. Lang, *Modular Units*, Grundlehren math. Wiss. **244**, Springer, New York, 1981.

[12] S. Lang, *Elliptic functions*, With an appendix by J. Tate. Second edition. Graduate Texts in Mathematics **112**. Springer-Verlag, New York, 1987 .

[13] J. P. SERRE, Propriétés galoisiennes des points d'ordre fini des courbes elliptiques, *Invent. Math.* **15** (1972), 259–331.

[14] J.-P. SERRE, *Représentations linéaires des groupes finis*, Hermann, 1967.

[15] J. P. SERRE, *Lectures on the Mordell-Weil Theorem*, 3rd edition, Vieweg & Sohn, Braunschweig/Wiesbaden, 1997

[16] G. SHIMURA, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan **11**, Kanô memorial lectures, **1**, Iwami Shoten Publishers and Princeton University Press, 1971.

[17] M. STRAMBI, Ph. D. Thesis, in preparation.