

JOURNAL

de Théorie des Nombres

de BORDEAUX

anciennement Séminaire de Théorie des Nombres de Bordeaux

François MAUCOURANT

Size of discriminants of periodic geodesics on the modular surface

Tome 37, n° 3 (2025), p. 795-835.

<https://doi.org/10.5802/jtnb.1343>

© Les auteurs, 2025.



Cet article est mis à disposition selon les termes de la licence
CREATIVE COMMONS ATTRIBUTION – PAS DE MODIFICATION 4.0 FRANCE.
<http://creativecommons.org/licenses/by-nd/4.0/fr/>



*Le Journal de Théorie des Nombres de Bordeaux est membre du
Centre Mersenne pour l'édition scientifique ouverte*

<http://www.centre-mersenne.org/>

e-ISSN : 2118-8572

Size of discriminants of periodic geodesics on the modular surface

par FRANÇOIS MAUCOURANT

RÉSUMÉ. Considérons une matrice aléatoire γ de $\Gamma = \mathrm{SL}(2, \mathbb{Z})$. Notons par $\mathcal{O}_{\mathbb{K}}$ l'anneau de Dedekind engendré par ses valeurs propres, et soient $\Delta_{\mathbb{K}}$, Δ_{γ} et $\Delta = \mathrm{Tr}(\gamma)^2 - 4$ les discriminants respectifs des anneaux $\mathcal{O}_{\mathbb{K}}$, de l'anneau des multiplicateurs $M(2, \mathbb{Z}) \cap \mathbb{Q}[\gamma]$ et $\mathbb{Z}[\gamma]$. Nous démontrons que les rapports entre ces discriminants convergent en loi vers une loi de probabilité lorsque les matrices de Γ sont ordonnées selon leur norme de Frobenius. En particulier, 42% des éléments de Γ ont un discriminant qui est fondamental, et $\mathbb{Z}[\gamma]$ est un anneau d'entier avec probabilité 32%.

ABSTRACT. Pick a random matrix γ in $\Gamma = \mathrm{SL}(2, \mathbb{Z})$. Denote by $\mathcal{O}_{\mathbb{K}}$ the Dedekind ring generated by its eigenvalues, and let $\Delta_{\mathbb{K}}$, Δ_{γ} and $\Delta = \mathrm{Tr}(\gamma)^2 - 4$ be the respective discriminant of the rings $\mathcal{O}_{\mathbb{K}}$, the multiplier ring $M(2, \mathbb{Z}) \cap \mathbb{Q}[\gamma]$ and $\mathbb{Z}[\gamma]$. We show that their ratios admit probability limit distributions when ordered by Frobenius norm. In particular, 42% of the elements of Γ have a fundamental discriminant, and $\mathbb{Z}[\gamma]$ is a ring of integers with probability 32%.

1. Introduction

1.1. Random geodesics. Let $\Gamma = \mathrm{SL}(2, \mathbb{Z})$. To a periodic oriented geodesic on the modular surface $\Gamma \backslash \mathbb{H}^2$ are associated several arithmetic quantities. The purpose of this article is to give an idea of their typical size. To pick a periodic geodesic at random, we choose a matrix $\gamma \in \Gamma$ and consider its conjugacy class $[\pm\gamma]$ in $\mathrm{PSL}(2, \mathbb{Z})$, as follows. For a large parameter $T > 0$, consider the ball of radius T in Γ :

$$\Gamma_T = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}) : \|\gamma\| \leq T\},$$

where $\|\cdot\|$ is the Frobenius norm $\|\gamma\| = \sqrt{\mathrm{Tr}({}^t\gamma\gamma)}$. A classical result is the asymptotic $|\Gamma_T| \sim 6T^2$ (see e.g. [21]). Denote by \mathbb{P}_T the uniform probability measure on Γ_T . Pick a random matrix γ using \mathbb{P}_T .

Manuscrit reçu le 27 mars 2024, révisé le 20 juillet 2025, accepté le 19 septembre 2025.

2020 *Mathematics Subject Classification.* 37A44, 11K65.

Mots-clefs. discriminant, random matrix.

This research was partially funded by the European Research Council (ERC GOAT 101053021).

We recall that γ is likely to be *hyperbolic*, more precisely for all $\eta > 0$,

$$(1.1) \quad \mathbb{P}_T(|\mathrm{Tr}(\gamma)| > 2) = 1 - O(T^{-1+\eta}).$$

It is also known that the trace $\mathrm{Tr}(\gamma)$ has roughly size T :

Proposition 1.1 (see Section 2.4). *For all $\alpha < \beta$,*

$$(1.2) \quad \lim_{T \rightarrow +\infty} \mathbb{P}_T\left(\frac{\mathrm{Tr}(\gamma)}{T} \in (\alpha, \beta)\right) = \frac{2}{\pi} \int_{\sup(\alpha, -1)}^{\inf(\beta, 1)} \sqrt{1-x^2} \, dx.$$

Assuming that γ is hyperbolic, its eigenvalues $\epsilon_\gamma, \epsilon_\gamma^{-1}$, where $|\epsilon_\gamma| > 1$, are units of the ring of integers $\mathcal{O}_{\mathbb{K}}$ of the real quadratic field $\mathbb{K} = \mathbb{Q}(\sqrt{\mathrm{Tr}(\gamma)^2 - 4})$. The unit ϵ_γ is close to $\mathrm{Tr}(\gamma)$, and ϵ_γ/T has the same asymptotic distribution (1.2) as $\mathrm{Tr}(\gamma)/T$. By a change of variable (see Section 2.4), this implies that in terms of the *length* $\ell_\gamma = 2 \log(|\epsilon_\gamma|)$ of the periodic geodesic associated to the conjugacy class $[\pm\gamma] \subset \mathrm{PSL}(2, \mathbb{Z})$, it is just a bit smaller than $2 \log(T)$: for all $\beta \geq 0$,

$$(1.3) \quad \lim_{T \rightarrow +\infty} \mathbb{P}_T(2 \log(T) - \beta \leq \ell_\gamma \leq 2 \log(T)) = \frac{2}{\pi} \int_0^\beta \sqrt{e^{-z} - e^{-2z}} \, dz.$$

Our first result is that the eigenvalue $|\epsilon_\gamma|$ is likely to be the *fundamental unit* $\epsilon_{\mathbb{K}}$ of $\mathcal{O}_{\mathbb{K}}$, so ℓ_γ is twice the *regulator* $R_{\mathbb{K}} = \log \epsilon_{\mathbb{K}}$ of the quadratic field \mathbb{K} . In particular, the regulator is essentially $\log(T) + O(1)$.

Theorem 1.2. *For all $\eta > 0$,*

$$(1.4) \quad \mathbb{P}_T(|\mathrm{Tr}(\gamma)| > 2, |\epsilon_\gamma| = \epsilon_{\mathbb{K}}) = 1 - O(T^{-\frac{1}{2}+\eta}).$$

In particular, since the norm of ϵ_γ is $\det(\gamma) = +1$, the ring $\mathcal{O}_{\mathbb{K}}$ is unlikely to have a unit of negative norm.

This property is a bit stronger than the trite fact that γ is likely to be primitive (i.e. not a power in Γ): it means that $\mathrm{sgn}(\mathrm{Tr}(\gamma))\gamma$ is not a power in $\mathrm{GL}(2, \mathbb{Q})$, and implies that the length of the associated geodesic is not a strict multiple of the length of another periodic geodesic (see Section 2.3). For example, primitive matrices like

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2, \quad \begin{pmatrix} 26 & 135 \\ 5 & 26 \end{pmatrix} = \begin{pmatrix} 2 & 9 \\ 1/3 & 2 \end{pmatrix}^3,$$

are quite rare. In the first example, the matrix is primitive in $\mathrm{SL}(2, \mathbb{Z})$ but is the square of a rational (here, integral) matrix of determinant -1 , indicating the presence of a unit of negative norm in $\mathcal{O}_{\mathbb{K}}$. In the second example, the length of the associated geodesic is 3 times the length of the geodesic given by $\pm(\frac{2}{3} \frac{1}{2})$.

1.2. Discriminants. We now introduce other arithmetical data attached to γ . The matrix ring $\mathbb{Q}[\gamma]$ is isomorphic to \mathbb{K} and it will be convenient to identify the ring of integers $\mathcal{O}_{\mathbb{K}}$ with the corresponding subring of $\mathbb{Q}[\gamma]$. The *multiplier ring* can be defined by

$$\mathcal{O}_{\gamma} = \mathbb{Q}[\gamma] \cap M(2, \mathbb{Z}).$$

The *discriminant* Δ_{γ} of the periodic geodesic corresponding to the conjugacy class of $\pm\gamma$ in $\mathrm{PSL}(2, \mathbb{Z})$ can be defined as the discriminant of the multiplier ring \mathcal{O}_{γ} . The link with more classical definitions using the quadratic form associated to the closed geodesic will be made in Section 2.2. Let u_{γ}, f_{γ} be the successive indexes in the inclusions

$$\mathbb{Z}[\gamma] \subset \mathcal{O}_{\gamma} \subset \mathcal{O}_{\mathbb{K}},$$

namely $u_{\gamma} = [\mathcal{O}_{\gamma} : \mathbb{Z}[\gamma]]$, $f_{\gamma} = [\mathcal{O}_{\mathbb{K}} : \mathcal{O}_{\gamma}]$. The number f_{γ} is called the *conductor* of the multiplier ring. It turns out that if we write $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $u_{\gamma} = \gcd(c, d - a, b)$ (see Lemma 2.1). Moreover,

$$\begin{cases} \mathrm{Tr}(\gamma)^2 - u_{\gamma}^2 \Delta_{\gamma} = 4, \\ f_{\gamma}^2 \Delta_{\mathbb{K}} = \Delta_{\gamma}, \end{cases}$$

where $\Delta_{\mathbb{K}}$ is the (fundamental) discriminant of \mathbb{K} . Notice that the first equation can be interpreted as a Pell Equation

$$X^2 - \Delta_{\mathbb{K}} Y^2 = 4,$$

and $(|\mathrm{Tr}(\gamma)|, u_{\gamma} f_{\gamma})$ is then, by Theorem 1.2, the fundamental solution of this equation with high probability. Our main result is that u_{γ}, f_{γ} are usually pretty small:

Theorem 1.3. *The joint distribution of (u_{γ}, f_{γ}) converges to a probability as $T \rightarrow +\infty$: there exist $(q_{n,m})_{n \geq 1, m \geq 1}$ positive numbers such that $\sum_{n,m \geq 1} q_{n,m} = 1$, with the property that for all $(n, m) \in \mathbb{N}^2$,*

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T(u_{\gamma} = n, f_{\gamma} = m) = q_{n,m}.$$

In fact, the probability $q_{n,m}$ will be shown to be a product of local factors

$$(1.5) \quad q_{n,m} = \prod_{p \text{ prime}} \mu_p(E_{p, v_p(n), v_p(m)}),$$

where $v_p(n), v_p(m)$ are the p -adic valuations of n, m , $E_{p, \alpha, \beta}$ some clopen subset in the group of p -adic matrices $\mathrm{SL}(2, \mathbb{Z}_p)$, and μ_p the normalized Haar measure on the latter group. The formulas for these local factors are given in Propositions 7.5 and A.4.

So, since u_{γ} and f_{γ} are usually small, both the discriminant Δ_{γ} and the fundamental discriminant $\Delta_{\mathbb{K}}$ are usually roughly of size T^2 . Some special values of n, m have interesting interpretation:

Corollary 1.4. *74% of the quadratic forms*

$$\tilde{Q}_\gamma := \operatorname{sgn}(\operatorname{Tr}(\gamma)) \det \left(\begin{pmatrix} X \\ Y \end{pmatrix}, \gamma \begin{pmatrix} X \\ Y \end{pmatrix} \right),$$

are primitive, that is

$$\begin{aligned} \lim_{T \rightarrow \infty} \mathbb{P}_T(\Delta_\gamma = \operatorname{Tr}(\gamma)^2 - 4) &= \sum_m q_{1,m} \\ &= \frac{5}{6} \prod_{p \text{ prime}, p > 2} \left(1 - \frac{2}{p(p^2 - 1)} \right) \simeq 0.7439 \dots \end{aligned}$$

We recover, partially and in a slightly different setting, a result of Raulf [17, Theorem 4.2]:

Corollary 1.5. *The discriminant Δ_γ is equal to the fundamental discriminant with probability 42%:*

$$\begin{aligned} \lim_{T \rightarrow \infty} \mathbb{P}_T(\Delta_\gamma = \Delta_{\mathbb{K}}) &= \sum_n q_{n,1} \\ &= \frac{75}{112} \prod_{p > 2} \left(1 - \frac{2p}{p^3 - 1} \right) \simeq 0.4269 \dots \end{aligned}$$

Raulf uses a different enumeration of geodesics, by length rather than Frobenius norm, and moreover she gives a remainder terms. But the density 0.4269.. is exactly the same (compare with [7, Section 4.3]). Another related paper is [3], where Bourgain and Kontorovich prove the existence of fundamental geodesics in thin semigroups, a significantly harder problem than the lattice case we are looking at here. Also, a related problem was recently considered by Bhargava, Shankar, and Wang, who also established a positive density result for square-free discriminants of quadratic forms [2, Theorem 1].

Corollary 1.6. *32% of matrices γ in Γ are such that $\mathbb{Z}[\gamma]$ is a Dedekind ring:*

$$\begin{aligned} \lim_{T \rightarrow \infty} \mathbb{P}_T(\mathbb{Z}[\gamma] \simeq \mathcal{O}_{\mathbb{K}}) &= q_{1,1} \\ &= \frac{7}{12} \prod_{p > 2} \left(1 - \frac{2(p^2 + p - 1)}{p^2(p^2 - 1)} \right) \simeq 0.3267 \dots \end{aligned}$$

This is analogous to the following result. For number fields of fixed degree n chosen by picking a random monic integer polynomial f of degree n , Bhargava, Shankar, and Wang [1, Theorem 1.2] established that $\mathbb{Z}[X]/(f)$ is the integer ring of the field with positive probability.

1.3. Class numbers. The last arithmetic quantity we want to mention are the *class numbers*. Recall that the narrow class number $h(\Delta)$ for a discriminant Δ of a (real) quadratic ring \mathcal{O} is the number of classes of invertible fractional ideals modulo products by element of positive norm. It happens to be also the number of periodic geodesics of $\Gamma \backslash \mathbb{H}^2$ of discriminant Δ . Geodesics with the same discriminant share the same multiplier ring (namely the unique real quadratic ring of the given discriminant), and length, although two geodesics of the same length may have different discriminants, for example the ones given by the conjugacy classes of $\pm \begin{pmatrix} 49 & 26 \\ 32 & 17 \end{pmatrix}$ and $\pm \begin{pmatrix} 33 & 136 \\ 8 & 33 \end{pmatrix}$, both primitive of length $2 \ln(33 + 8\sqrt{17})$ and of respective discriminants 1088 and 68. The average of class numbers according to length of the geodesics is studied in [7, 18, 18, 22]. From Theorem 1.3 and the Brauer-Siegel Theorem, we can deduce that those class numbers are typically of size $T^{1 \pm \epsilon}$:

Corollary 1.7. *The quantity $\log h(\Delta_{\mathbb{K}}) / \log T$ converges in probability to 1, that is for all $\eta > 0$,*

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T \left(\left| \frac{\log h(\Delta_{\mathbb{K}})}{\log T} - 1 \right| > \eta \right) = 0.$$

Similarly, the class number $h(\Delta_{\gamma})$ satisfies for all $\eta > 0$,

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T \left(\left| \frac{\log h(\Delta_{\gamma})}{\log T} - 1 \right| > \eta \right) = 0.$$

In the case of enumeration of geodesics by length, the above corollary follows from a much stronger result of Raulf [18, Theorem A].

1.4. Concluding remarks. It must be noted that the probability \mathbb{P}_T used here is in a sense “not far” from the uniform probability of the set of periodic geodesics of length $\leq 2 \log(T)$. It may be surprising because this way to pick a geodesic is very redundant: there are many representatives γ of the same conjugacy class $[\pm \gamma]$. But if one restricts to subsets of Γ that are invariant by conjugacy, hyperbolic and invariant under $\gamma \mapsto -\gamma$ (and hence correspond to sets of oriented periodic geodesics), the two family of probabilities share the same sets of positive upper density. I don’t know a reference for this folklore fact, so a statement and its proof are included in Appendix B.

Actually, it seems reasonable to expect that for many sets of periodic geodesics, the densities should be the same for the enumeration of geodesics by length, or when using the probability \mathbb{P}_T . An example where this equality happens is given by Corollary 1.5 and the result of Raulf ([17, Theorem 4.2]), namely the density of fundamental discriminants. Another example is equality is a result of Sarnak [22, Theorem 3.4] about the density of discriminants for which some fixed prime p divides the number u_{γ} .

Raulf [17, 18] and Hashimoto [7] studied the density of discriminants in a given arithmetic progression. Although there is a link with Theorem 2, it does not seem obvious to relate those densities with the number $q_{n,m}$.

If picking a matrix at random using \mathbb{P}_T is a bit less natural than the ordering of geodesics by length, it has the advantage of being very easy to implement: we refer the reader to [19] for this. The algorithm described there by Rivin follows a probability slightly different from \mathbb{P}_T (but close enough), and is much more efficient than the naive enumeration of all matrices of norm $\leq T$ (or of all primitive quadratic forms of given discriminant). In particular, the numbers in Corollaries 1.4, 1.5 and 1.6 are consistent with numerical experiments.

This way to pick a random geodesic gives very large discriminants and class number when compared to the size of the fundamental unit. This phenomenon was called the “*discriminant-regulator paradox*” by McMullen in [11, Chapter 28]. The explanation is that number theorists usually order quadratic fields by discriminant, while from the geometrical point of view, the fields appear ordered by the length of geodesics (which is twice regulator). To quote Sarnak [23],

“That the class numbers are large on average when ordered by the size of the unit is not surprising, as we are favouring discriminants with small units. . .”

Another interesting arithmetical quantity associated to a periodic geodesic is the Rademacher invariant. Its behaviour was analyzed by Mozzochi in [13].

1.5. Plan of the paper. In Section 2, we discuss the arithmetic objects attached to a hyperbolic matrix γ . In Section 3, we prove Theorem 1.2: the largest eigenvalue of γ is the fundamental unit with probability close to 1.

In Section 4, we discuss the p -adic formalism and uniform equidistribution results modulo congruence subgroups that will help to establish Theorem 1.3.

In Section 5, we give the strategy to prove Theorem 1.3, reducing it to three facts, the main two being that knowledge of the p -adic valuation of the indexes u_γ, f_γ can be determined by looking at the reduction of the matrix γ modulo p^k , for a suitable k , and that the discriminant typically does not have large prime square factors.

In Section 6, we establish the first claim about the p -adic valuation of the conductors. In Section 7, we determine the local factors $\mu_p(E_{p,\alpha,\beta})$, by counting matrices in $\mathrm{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$ in an elementary fashion. Note that in these two sections, in order to keep things simple, the prime number p is usually assumed to be $\neq 2$, while the somewhat tedious case $p = 2$ is postponed to Appendix A.

In Section 8, we establish the crucial fact that most discriminants are not divisible by the square of a large prime (Proposition 8.1), using the uniform counting on congruence subgroups.

In Section 9, we explain how to get Corollary 1.7.

In Appendix A, we detail the computations of the local factors for the exceptional prime $p = 2$. In Appendix B, we prove that enumerating matrices by norm, or periodic geodesic by length, yield the same negligible sets, for properties which are conjugacy invariants of $\mathrm{PSL}(2, \mathbb{Z})$.

2. Arithmetic quantities associated to a matrix γ

Let us now review in more detail the arithmetic data we are interested in.

2.1. Discriminant. Let $\gamma \in \Gamma$ be a matrix

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Its characteristic polynomial is:

$$\chi_\gamma(X) = \det(\gamma - XI_2) = X^2 - \mathrm{Tr}(\gamma)X + 1.$$

We will assume in what follows that γ is hyperbolic, that is $|\mathrm{Tr}(\gamma)| > 2$. In this case the polynomial χ_γ is irreducible in $\mathbb{Q}[X]$, the field $\mathbb{K} = \mathbb{Q}[\gamma] \simeq \mathbb{Q}[X]/(\chi_\gamma) \simeq \mathbb{Q}(\sqrt{\mathrm{Tr}(\gamma)^2 - 4})$ containing the eigenvalues of γ is a real quadratic field. The positive integer $\mathrm{Tr}(\gamma)^2 - 4$ might not be square-free, but we can always find the unique integers $m > 0$ and square-free $D > 1$ such that

$$\mathrm{Tr}(\gamma)^2 - 4 = m^2 D.$$

Then the ring of integers of \mathbb{K} is

$$\mathcal{O}_{\mathbb{K}} = \begin{cases} \mathbb{Z}[\sqrt{D}] & \text{if } D \equiv 2, 3 \pmod{4}, \\ \mathbb{Z}\left[\frac{1+\sqrt{D}}{2}\right] & \text{if } D \equiv 1 \pmod{4}. \end{cases}$$

Its discriminant is fundamental and equal to

$$\Delta_{\mathbb{K}} = \begin{cases} 4D & \text{if } D \equiv 2, 3 \pmod{4}, \\ D & \text{if } D \equiv 1 \pmod{4}. \end{cases}$$

We can also consider the ring $\mathbb{Z}[\gamma] \simeq \mathbb{Z}[X]/(\chi_\gamma)$. The discriminant of this ring is equal to $\mathrm{Tr}(\gamma)^2 - 4$. We define the multiplier ring \mathcal{O}_γ as the stabilizer of \mathbb{Z}^2 :

$$\mathcal{O}_\gamma = \{\gamma' \in \mathbb{Q}[\gamma] : \gamma'(\mathbb{Z}^2) \subset \mathbb{Z}^2\} = \mathbb{Q}[\gamma] \cap M(2, \mathbb{Z}).$$

Recall that we defined the discriminant Δ_γ as the discriminant of this ring. All orders, i.e. subrings of \mathbb{K} which are \mathbb{Z} -modules of rank 2, are of the form $\mathbb{Z}[f\sqrt{D}]$ or $\mathbb{Z}[f\frac{1+\sqrt{D}}{2}]$ depending if $D \equiv 2, 3$, or $1 \pmod{4}$, for some integer f called the conductor of the ring. The conductor is also the index of the ring

in the maximal order $\mathcal{O}_{\mathbb{K}}$, and its discriminant is $f^2\Delta_{\mathbb{K}}$. Here we defined f_{γ} as the conductor of the multiplier ring \mathcal{O}_{γ} , so

$$\Delta_{\gamma} = f_{\gamma}^2 \Delta_K.$$

By definition, u_{γ} is the index of $\mathbb{Z}[\gamma]$ in \mathcal{O}_{γ} , so

$$\mathrm{Tr}(\gamma)^2 - 4 = u_{\gamma}^2 \Delta_{\gamma},$$

and m is either $u_{\gamma}f_{\gamma}$ or $2u_{\gamma}f_{\gamma}$ depending on $D \bmod 4$.

From a practical point of view, given the coefficients a, b, c, d of γ , m and D are deduced from the trace by factoring possible squares factors of $(a+d)^2 - 4$, $\Delta_{\mathbb{K}}$ is computed using D , but it remains to identify u_{γ} and f_{γ} as factors of m . There is indeed a simple formula for u_{γ} :

Lemma 2.1. *The index of $\mathbb{Z}[\gamma]$ in \mathcal{O}_{γ} is given by:*

$$u_{\gamma} = \gcd(c, d - a, b)$$

Proof. The set \mathcal{O}_{γ} is defined by

$$\mathcal{O}_{\gamma} = M(2, \mathbb{Z}) \cap \{\lambda I_2 + \mu \gamma : (\lambda, \mu) \in \mathbb{Q}^2\}.$$

If we set $u' = \gcd(c, d - a, b)$, we claim that it is equal to

$$\mathcal{O}' := \mathbb{Z}I_2 + \mathbb{Z}\frac{\gamma - aI_2}{u'}.$$

This ring obviously has $\mathbb{Z}[\gamma]$ as a subring of index u' , so the equality $\mathcal{O}_{\gamma} = \mathcal{O}'$ is sufficient to conclude that $u_{\gamma} = u'$. By definition of u' , $\frac{\gamma - aI_2}{u'}$ has indeed integer coefficients, so $\mathcal{O}_{\gamma} \subset \mathcal{O}'$. Reciprocally, let $\lambda, \mu \in \mathbb{Q}^2$, such that $\gamma' = \lambda I_2 + \mu \gamma$ has integer coefficients, and write $\mu = \frac{v}{w}$, with v, w integers, $w > 0$ and coprime to v . Then

$$\gamma' = \begin{pmatrix} \lambda + \frac{va}{w} & \frac{vb}{w} \\ \frac{vc}{w} & \lambda + \frac{vd}{w} \end{pmatrix},$$

but since $\lambda + \frac{va}{w}$ is an integer, say n , the matrix

$$\gamma' - nI_2 = \begin{pmatrix} 0 & \frac{vb}{w} \\ \frac{vc}{w} & \frac{v(d-a)}{w} \end{pmatrix} = \frac{v(\gamma - aI_2)}{w},$$

is also integral, so w divides $vc, vb, v(d-a)$, and since v is coprime to w , w divides u' , so $\gamma' - nI_2 \in \mathbb{Z}\frac{(\gamma - aI_2)}{u'}$, and $\gamma' \in \mathcal{O}'$ as requested. \square

2.2. Quadratic forms. Although we will not use this, here we make the link with the more classical definition of the discriminant Δ_{γ} using quadratic forms. The quadratic form

$$\tilde{Q}_{\gamma}(X, Y) := \mathrm{sgn}(\mathrm{Tr}(\gamma)) \det\left(\begin{pmatrix} X \\ Y \end{pmatrix}, \gamma \begin{pmatrix} X \\ Y \end{pmatrix}\right),$$

which is obviously γ -invariant with integer coefficients, can be expressed

$$\tilde{Q}_\gamma(X, Y) = \operatorname{sgn}(\operatorname{Tr}(\gamma))(cX^2 + (d - a)XY - bY^2).$$

Its discriminant (as a quadratic form) is, again, $\operatorname{Tr}(\gamma)^2 - 4$. The factor $\operatorname{sgn}(\operatorname{Tr}(\gamma))$ in the expression is here to insure that $\tilde{Q}_\gamma = \tilde{Q}_{-\gamma}$, so the $\operatorname{SL}(2, \mathbb{Z})$ -class of \tilde{Q}_γ depends only on the class of γ in $\operatorname{PSL}(2, \mathbb{Z})$.

The form \tilde{Q}_γ is not necessarily primitive, meaning the gcd of its coefficients (here u_γ by Lemma 2.1), may not be 1. So generally one consider instead the quadratic form

$$Q_\gamma = \frac{\tilde{Q}_\gamma}{u_\gamma},$$

see for example [3, 5, 8, 22]. The discriminant of this quadratic form is then Δ_γ , which is the usual definition for the discriminant of an hyperbolic conjugacy class of $\operatorname{PSL}(2, \mathbb{Z})$. The discriminant characterize the multiplier ring, that is two primitive hyperbolic matrices have the same discriminant if and only if their multiplier rings are isomorphic. This situation is specific to the quadratic case. The set of geodesics sharing the same discriminant Δ can be identified with the set of $\operatorname{SL}(2, \mathbb{Z})$ -classes of primitives integer quadratic forms of discriminant Δ . This set also identifies with the narrow class group of their common multiplier ring, that is the group of invertible fractional ideals modulo multiplication by elements of positive norm. The cardinal of this set is called the *class number* $h(\Delta)$.

2.3. Units. We denote by

$$\{\epsilon_\gamma, \bar{\epsilon}_\gamma\} = \left\{ \frac{\operatorname{Tr}(\gamma) \pm \sqrt{\operatorname{Tr}(\gamma)^2 - 4}}{2} \right\},$$

the two eigenvalues of γ , chosen so that $|\epsilon_\gamma| > 1 > |\bar{\epsilon}_\gamma|$. The number $|\epsilon_\gamma|$ is thus a positive unit in the field \mathbb{K} . The quadratic field \mathbb{K} has a *fundamental unit* $\epsilon_{\mathbb{K}} \in \mathcal{O}_{\mathbb{K}}$, a *regulator* $R_{\mathbb{K}} = \log \epsilon_{\mathbb{K}}$, and a *class number* $h(\Delta_{\mathbb{K}})$. The norm of the fundamental unit $N_{\mathbb{K}/\mathbb{Q}}(\epsilon_{\mathbb{K}}) = \epsilon_{\mathbb{K}}\bar{\epsilon}_{\mathbb{K}}$ is either ± 1 . By Dirichlet units' Theorem, there exists an integer e_γ such that $|\epsilon_\gamma| = (\epsilon_{\mathbb{K}})^{e_\gamma}$. Recall that the field $\mathbb{Q}[\gamma]$ can be identified with \mathbb{K} , and in this case the matrix trace and determinant correspond to the algebraic trace and norm respectively. Therefore, the fundamental unit can be identified with a rational matrix (since it is in $\mathbb{Q}[\gamma]$), and $\epsilon_K = |\epsilon_\gamma|$ if and only if $\operatorname{sgn}(\operatorname{Tr}(\gamma))\gamma$ is a primitive matrix in $\operatorname{GL}(2, \mathbb{Q})$.

2.4. Distribution of the trace. Here we indicate how to recover the formula in Proposition 1.1 and the asymptotic (1.3). It may be worth mentioning that these two formulas depends on the choice of the norm $\|\cdot\|$ considered in the definition of Γ_T , here the Frobenius norm $\|\gamma\| = \sqrt{\operatorname{Tr}({}^t\gamma\gamma)}$, while Theorems 1.2 and 1.3 do not.

By [10, Theorem 2], applied to the characteristic function $\chi_{\alpha,\beta}$ of the set of matrices of Frobenius norm ≤ 1 and of trace between $\alpha < \beta$,

$$\begin{aligned} \lim_{T \rightarrow +\infty} \mathbb{P}_T \left(\frac{\text{Tr}(\gamma)}{T} \in (\alpha, \beta) \right) \\ = \frac{1}{2\pi^2} \int_0^1 \int_0^{2\pi} \int_0^{2\pi} \chi_{\alpha,\beta} \begin{pmatrix} r \sin \theta \sin \varphi & r \cos \theta \sin \varphi \\ r \sin \theta \cos \varphi & r \cos \theta \cos \varphi \end{pmatrix} r \, dr \, d\theta \, d\varphi. \end{aligned}$$

Note that [10, Theorem 2] assumes the function to be continuous, and $\chi_{\alpha,\beta}$ is not. Here $\chi_{\alpha,\beta}$ can be approximated from above or below by continuous functions having approximately the same integral with respect to the limit measure, so this formula is also valid in this case. Since the trace of the above matrix is $r \sin(\theta + \varphi)$, the above integral is more easily computed using the change of variables

$$(r, \theta, \varphi) \mapsto (x = r \sin(\theta + \varphi), y = r \cos(\theta + \varphi), \theta),$$

which sends $(0, 1) \times (0, 2\pi) \times (0, 2\pi)$ to $\mathbb{D} \times [0, 2\pi]$, the set of matrices of trace in (α, β) into $\{x \in (\alpha, \beta), (x, y) \in \mathbb{D}, \theta \in (0, 2\pi)\}$ and changes the polar coordinate measure $r \, dr \, d\theta \, d\varphi$ into $dx \, dy \, d\theta$. We thus get

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T \left(\frac{\text{Tr}(\gamma)}{T} \in (\alpha, \beta) \right) = \frac{2}{\pi} \int_{\sup(\alpha, -1)}^{\inf(\beta, 1)} \sqrt{1 - x^2} \, dx,$$

as claimed.

To recover the asymptotic (1.3), we first notice that provided that γ is hyperbolic, $|\epsilon_\gamma - \text{Tr}(\gamma)| \leq 1$, so Proposition 1.1 holds for $\frac{\epsilon_\gamma}{T}$ instead of $\frac{\text{Tr}(\gamma)}{T}$. Recall that the length is given by $\ell_\gamma = 2 \log(|\epsilon_\gamma|)$. Then (1.3) is obtained from the previous formula by the change of variable $z = -2 \log(|x|)$.

3. Eigenvalues are fundamental units

In this section, we prove Theorem 1.2, which states that with high probability, the absolute value of the eigenvalue ϵ_γ of the randomly picked matrix γ is the fundamental unit of the integer ring of the underlying field $\mathbb{Q}[\gamma]$. The idea is that if that does not happen, then the trace of γ must lie in the integer image of a family of polynomials P_k of degree $k \geq 2$, which is a fairly scarce set in \mathbb{N} . We first prove three useful lemmata.

3.1. Almost all matrices are hyperbolic. The following Lemma is a minor variant of [3, Lemma 7.1].

Lemma 3.1. *For all $\eta > 0$, there exists $C_\eta > 0$ such that for all $t \in \mathbb{Z}$ and all $T \geq 2$,*

$$|\{\gamma \in \Gamma_T : \text{Tr}(\gamma) = t\}| \leq C_\eta T^{1+\eta}.$$

Proof. We can and will assume that $|t| \leq 2T$, otherwise the set $\{\gamma \in \Gamma_T : \text{Tr}(\gamma) = t\}$ is empty so the bound is trivial.

Let $\gamma = \begin{pmatrix} a & b \\ c & t-a \end{pmatrix}$ be an integer matrix of norm $\leq T$, trace t and determinant 1. There is at most possible $2T + 1$ choices for a in $[-T, T]$. Since the determinant is one, we have $bc = -a^2 + at - 1$, so for each a , the integer b must be chosen among the divisors of $|-a^2 + at - 1| \leq 3T^2 + 1 \leq 4T^2$. Recall (see e.g. [6, Theorem 315]) that for all $\eta > 0$, there exists a $c_\eta > 0$ such that the number of divisors $d(N)$ of an integer $N \geq 1$ satisfies $d(N) \leq c_\eta N^{\eta/2}$. So we have $2T + 1$ choices for a , given a we have at most $c_\eta (4T^2)^{\eta/2}$ choices for b , and c is deduced from a, b, t . Overall, there is at most $O(T^{1+\eta})$ possible choices, where the implied constant depends only on η . \square

When applied to the finite set of values $t \in \{-2, -1, 0, 1, 2\}$, and recalling the well-known asymptotic for $|\Gamma_T|$:

$$|\Gamma_T| \sim 6T^2,$$

(see for example [21, Theorem 1.1]), we recover the claim (1.1) of the introduction:

Corollary 3.2. *For any $\eta > 0$,*

$$\mathbb{P}_T(|\text{Tr}(\gamma)| > 2) = 1 - O_\eta(T^{-1+\eta}).$$

3.2. Traces of powers.

Lemma 3.3. *Let $\gamma \in \text{GL}(2, \mathbb{C})$ be a matrix such that $\det(\gamma) \in \{\pm 1\}$. Then*

$$\text{Tr}(\gamma^2) = \text{Tr}(\gamma)^2 \mp 2.$$

Proof. This follows immediately for the characteristic equation

$$\gamma^2 - \text{Tr}(\gamma)\gamma + \det(\gamma)I_2 = 0. \quad \square$$

The trace of a power is in general a polynomial of the trace:

Lemma 3.4. *There exists a sequence $(P_k)_{k \geq 2}$ of polynomials, where P_k is of degree k such that for all matrices $\gamma \in \text{SL}(2, \mathbb{C})$,*

$$\text{Tr}(\gamma^k) = P_k(\text{Tr}(\gamma)),$$

and we have for $x > 2$,

$$P_k(x) \geq (x - 1)^k.$$

Moreover, if k is odd and $|x| > 2$, then x and $P_k(x)$ share the same sign. In fact, $P_k(X) = 2T_k(X/2)$, where $(T_k)_{k \geq 0}$ is the Chebyshev polynomial of the first kind.

Proof. Define the sequence of polynomials $(Q_k)_{k \geq 0}$ of degree k by the following recursion:

$$\begin{cases} Q_0(X) = 1, \\ Q_1(X) = X, \\ Q_{k+2}(X) = XQ_{k+1} - Q_k, \quad k \geq 0. \end{cases}$$

In other words, $Q_k(2X)$ are the Chebyshev polynomials of the second kind. If $\gamma \in \mathrm{SL}(2, \mathbb{C})$, we claim that

$$(3.1) \quad \forall k \geq 2, \quad \gamma^k = Q_{k-1}(\mathrm{Tr}(\gamma))\gamma - Q_{k-2}(\mathrm{Tr}(\gamma))I_2.$$

Indeed, writing $t = \mathrm{Tr}(\gamma)$, this is true for $k = 2$ by the characteristic equation

$$\gamma^2 = t\gamma - I_2,$$

and if for some $k \geq 2$,

$$\gamma^k = Q_{k-1}(t)\gamma - Q_{k-2}(t)I_2,$$

then, multiplying by γ and using the characteristic equation,

$$\begin{aligned} \gamma^{k+1} &= Q_{k-1}(t)(t\gamma - I_2) - Q_{k-2}(t)\gamma, \\ &= (tQ_{k-1}(t) - Q_{k-2}(t))\gamma - Q_{k-1}(t)I_2, \\ &= Q_k(t)\gamma - Q_{k-1}(t)I_2. \end{aligned}$$

Thus Equation (3.1) follows by induction. Now we define for $k \geq 2$ the polynomial P_k

$$P_k = XQ_{k-1} - 2Q_{k-2} = Q_k - Q_{k-2},$$

and it can be checked that $P_k(2X)/2$ agree with the Chebyshev polynomials of the first kind. By Equation (3.1),

$$\mathrm{Tr}(\gamma^k) = P_k(\mathrm{Tr}(\gamma)).$$

Let's now prove the lower bound $P_k(x) \geq (x-1)^k$ for all $x > 2$. Let $x > 2$, there exists an hyperbolic matrix γ in $\mathrm{SL}(2, \mathbb{C})$ whose trace is $x = \mathrm{Tr}(\gamma)$. Let λ, λ^{-1} be the two eigenvalues of γ , inverse of each other since $\det(\gamma) = 1$. Up to switching λ and its inverse, we can assume that $\lambda > 1 > \lambda^{-1} > 0$, so in particular

$$\lambda = \mathrm{Tr}(\gamma) - \lambda^{-1} > \mathrm{Tr}(\gamma) - 1 = x - 1,$$

so for any $k \geq 2$,

$$P_k(x) = \mathrm{Tr}(\gamma^k) = \lambda^k + \lambda^{-k} > (x-1)^k.$$

With the same notations, notice that λ, λ^{-1} have the same sign than $\mathrm{Tr}(\gamma)$ and $\mathrm{Tr}(\gamma^k)$ provided k is odd, so the statement about the sign of x and $P_k(x)$ follows. \square

3.3. Proof of Theorem 1.2. Let $T > 2$ be large. Consider the sets of matrices of norm less than T such that their trace is in the image the polynomials $X^2 + 2$ and P_k for some integer $\leq 2\sqrt{T}$:

$$F_{2,-}^T = \left\{ \gamma \in \Gamma_T : \exists t \in [1, 2\sqrt{T}] \cap \mathbb{Z}, \operatorname{Tr}(\gamma) = t^2 + 2 \right\},$$

$$\text{for } k \geq 2, \quad F_k^T = \left\{ \gamma \in \Gamma_T : \exists t \in [3, 2\sqrt{T}] \cap \mathbb{Z}, \operatorname{Tr}(\gamma) = P_k(t) \right\}.$$

Notice that:

Proposition 3.5. *Let $\gamma \in \Gamma$, hyperbolic, such that $|\epsilon_\gamma| \neq \epsilon_{\mathbb{K}}$ and $\|\gamma\| \leq T$. Then γ lies in the set*

$$F_T = \pm \left(F_{2,-}^T \cup \bigcup_{2 \leq k \leq 1 + \frac{\log T}{\log 2}} F_k^T \right).$$

In particular, we have

$$\mathbb{P}_T(|\operatorname{Tr}(\gamma)| > 2, |\epsilon_\gamma| \neq \epsilon_{\mathbb{K}}) \leq \frac{|F_T|}{|\Gamma_T|}$$

Proof. Up to considering $-\gamma$ instead of γ , we can assume that $\operatorname{Tr}(\gamma) > 2$.

First consider the case where the integer ring of $\mathbb{K} = \mathbb{Q}[\gamma]$ has a unit with negative norm. So ϵ_γ is necessarily an even power of the fundamental unit, so is a square : there exists $A \in \mathbb{Q}[\gamma]$ in the ring of integers of \mathbb{K} with determinant ± 1 , such that $A^2 = \gamma$. Its trace is an integer since A lies in the ring of integers. We can replace A by $-A$ to insure that $t := \operatorname{Tr}(A)$ is nonnegative. By Lemma 3.3, $\operatorname{Tr}(\gamma) = t^2 + 2$. Since $\operatorname{Tr}(\gamma) \leq 2T$, this implies that $t \leq \sqrt{2T - 2} \leq 2\sqrt{T}$, so $\gamma \in F_{2,-}^T$.

Now assume that the fundamental unit $\epsilon_{\mathbb{K}}$ has norm 1. Since by assumption $\epsilon_\gamma \neq \epsilon_{\mathbb{K}}$, then ϵ_γ is a k -th power of another unit of the integer ring of $\mathbb{Q}[\gamma]$, for some $k \geq 2$. So there exists $A \in \mathbb{Q}[\gamma]$ with integer trace $t = \operatorname{Tr}(A)$ and determinant 1 such that $A^k = \gamma$, so that

$$\operatorname{Tr}(\gamma) = P_k(t).$$

If k is even, one can replace A by $-A$ and assume that $t \geq 0$. If k is odd, t and $\operatorname{Tr}(\gamma)$ have the same sign so $t \geq 0$ also. In any case, since the matrix γ is hyperbolic, so is A and thus $t > 2$, so $t \geq 3$. For a upper bound on t , using Lemma 3.4:

$$2T \geq \operatorname{Tr}(\gamma) = P_k(t) \geq (t - 1)^k,$$

so

$$3 \leq t \leq \sqrt[k]{2T} + 1,$$

which we can bound using the crude estimate $\sqrt[k]{2T} + 1 \leq 2\sqrt{T}$ for $T > 2$ and $k \geq 2$. Now $T \geq \frac{1}{2}(t-1)^k \geq 2^{k-1}$ so

$$2 \leq k \leq 1 + \frac{\log(T)}{\log(2)}.$$

so $\gamma \in F_k^T$ with the above bound on k . □

We conclude the proof of Theorem 1.2 by bounding the number of elements in F_T . Elements of F_k^T have at most $2\sqrt{T}$ different traces. By Lemma 3.1, we have for every $\eta > 0$,

$$|F_k^T| \leq C_\eta(2\sqrt{T})T^{1+\eta} = 2C_\eta T^{3/2+\eta}.$$

Similarly,

$$|F_{2,-}^T| \leq C_\eta(2\sqrt{T})T^{1+\eta} = 2C_\eta T^{3/2+\eta}.$$

Adding these $2 + \log_2(T)$ upper bounds, we get

$$|F_T| \leq 2(2 + \log_2(T))C_\eta T^{3/2+\eta}$$

so, decreasing slightly the value of the parameter η in order to dismiss the logarithmic term, we get for T large enough,

$$\mathbb{P}_T(|\mathrm{Tr}(\gamma)| > 2, |\epsilon_\gamma| \neq \epsilon_{\mathbb{K}}) \leq \frac{|F_T|}{|\Gamma_T|} = O_\eta(T^{-\frac{1}{2}+\eta}).$$

Together with Corollary 3.2, this concludes the proof of Theorem 1.2.

4. Equidistribution Theorem, and p -adic formalism

4.1. The Uniform Equidistribution Theorem. Let $\Gamma(N)$ be the congruence subgroup of level N , that is the kernel of the surjective reduction map $\pi_N : \mathrm{SL}(2, \mathbb{Z}) \rightarrow \mathrm{SL}(2, \mathbb{Z}/N\mathbb{Z})$. It is classical that elements of Γ_T distribute evenly in each class mod $\Gamma(N)$: for any subset $E \subset \mathrm{SL}(2, \mathbb{Z}/N\mathbb{Z})$,

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T(\pi_N(\gamma) \in E) = \frac{|E|}{[\Gamma : \Gamma(N)]}.$$

We will need the following refinement, due to Nevo and Sarnak. They showed, using the uniform spectral gap for congruence subgroups, that

Theorem 4.1. [15, Theorem 3.2] *There exist constants $C > 0$, $\beta > 0$, $T_0 > 0$ such that for all integer $N \geq 2$, and all $\gamma_0 \in \Gamma$ and $T \geq T_0$,*

$$\left| |\gamma_0 \Gamma(N) \cap \Gamma_T| - \frac{|\Gamma_T|}{[\Gamma : \Gamma(N)]} \right| \leq C |\Gamma_T|^{1-\beta}.$$

4.2. The p -adic formalism. It will be convenient to rephrase the equidistribution Theorem (without the remainder terms given by Nevo and Sarnak) using the vocabulary of p -adic and profinite integers. This will allow to avoid reference to the level N that will be implicitly considered. When p is prime, we denote by \mathbb{Z}_p the ring of p -adic integers, and $\widehat{\mathbb{Z}} \simeq \prod_p \mathbb{Z}_p$ the ring of profinite integers. Let μ_p be the Haar measure on $\mathrm{SL}(2, \mathbb{Z}_p)$, normalized as a probability, and $\mu = \otimes_p \mu_p$ the Haar measure on $\mathrm{SL}(2, \widehat{\mathbb{Z}}) \simeq \prod_p \mathrm{SL}(2, \mathbb{Z}_p)$, by virtue of the Chinese Remainder Theorem. Recall that a clopen set (closed and open set) of this profinite group is nothing else than the preimage of a set by the reduction mod N , for some N depending on the clopen set. The set Γ_T can be seen as a subset of $\mathrm{SL}(2, \widehat{\mathbb{Z}})$ using the diagonal embedding $\mathrm{SL}(2, \mathbb{Z}) \hookrightarrow \mathrm{SL}(2, \widehat{\mathbb{Z}})$, and the equidistribution Theorem can be rewritten:

Theorem 4.2. (*Equidistribution Theorem mod N , rephrased*) For any clopen set $E \subset \mathrm{SL}(2, \widehat{\mathbb{Z}})$, we have

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T(\gamma \in E) = \mu(E).$$

5. Proof of Theorem 1.3

In this section, we prove Theorem 1.3, assuming three facts that will be proved later.

5.1. The three facts. These are:

- (1) When p is a prime number, the p -adic valuations $v_p(u_\gamma), v_p(f_\gamma)$ can be determined by looking at the reduction of γ in $\mathrm{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$ for sufficiently large $k \geq 1$. More precisely, there is a partition of the p -adic matrices of trace $\neq \pm 2$ into non-empty clopen sets:

$$\mathrm{SL}(2, \mathbb{Z}_p) - \mathrm{Tr}^{-1}(\{\pm 2\}) = \bigsqcup_{\alpha, \beta \geq 0} E_{p, \alpha, \beta},$$

such that for any hyperbolic matrix γ ,

$$(5.1) \quad (v_p(u_\gamma) = \alpha, v_p(f_\gamma) = \beta) \iff \gamma \in E_{p, \alpha, \beta}.$$

Moreover, the set of matrices of trace ± 2 in $\mathrm{SL}(2, \mathbb{Z}_p)$ has zero μ_p -measure.

- (2) The following product over all primes is convergent:

$$\prod_p \mu_p(E_{p, 0, 0}) = \frac{7}{12} \prod_{p \geq 3} \left(1 - \frac{2(p^2 + p - 1)}{p^2(p^2 - 1)} \right) > 0.$$

- (3) For any given $\epsilon > 0$, there exists $K(\epsilon) > 0$ such that with probability $1 - \epsilon$, the product $u_\gamma f_\gamma$ does not have any large prime factors $> K(\epsilon)$.

This means that for T large enough,

$$\mathbb{P}_T \left(u_\gamma = \prod_{p \leq K(\epsilon)} p^{v_p(u_\gamma)}, f_\gamma = \prod_{p \leq K(\epsilon)} p^{v_p(f_\gamma)} \right) \geq 1 - \epsilon.$$

These three facts and the Chinese Remainder Theorem imply that to recover u_γ and f_γ , it is (with probability close to 1) sufficient to look at the reduction mod N , for a suitable N . The first and second facts are easy and treated in Sections 6 and 7, but involve some lengthy elementary counting to get the formulas for the local factors $\mu_p(E_{p,\alpha,\beta})$, especially in the case $p = 2$ which is relegated to the appendix. The third one is the crucial step, and uses the spectral gap for congruence subgroups as expressed in Nevo and Sarnak's Theorem [15], and insures the tightness of the sequence of law of (u_γ, f_γ) as $T \rightarrow +\infty$. This will be done in Section 8.

5.2. Proof of Theorem 1.3. Assuming the three above facts, we now prove the Theorem. Given n, m positive integers, our goal is to show that for some $q_{n,m} > 0$ and for all $\epsilon > 0$, when T is large enough,

$$|\mathbb{P}_T(u_\gamma = n, f_\gamma = m) - q_{n,m}| \leq 3\epsilon.$$

Let K be larger than any prime factor of nm , such that $K \geq K(\epsilon)$ and moreover

$$(5.2) \quad \left| \prod_{p \leq K} \mu_p(E_{p,v_p(n),v_p(m)}) - \prod_p \mu_p(E_{p,v_p(n),v_p(m)}) \right| \leq \epsilon.$$

This is possible because the probabilities $\mu_p(E_{p,\alpha,\beta})$ are between 0 and 1, so the finite product decreases to its limit as $K \rightarrow \infty$. Moreover, the non-divergence to zero of the product $\prod_p \mu_p(E_{p,0,0})$ tells us that the limit

$$q_{n,m} := \prod_p \mu_p(E_{p,v_p(n),v_p(m)}),$$

is positive, as long as all the individual factors are nonzero. This latter property will be obvious from their explicit computation in Propositions 7.5 and A.4.

Let

$$G_{n,m,K} = \left(\prod_{p \leq K} E_{p,v_p(n),v_p(m)} \right) \times \prod_{p > K} \mathrm{SL}(2, \mathbb{Z}_p) \subset \mathrm{SL}(2, \widehat{\mathbb{Z}}).$$

This is a clopen set, so by the equidistribution Theorem 4.2,

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T(\gamma \in G_{n,m,K}) = \mu(G_{n,m,K}) = \prod_{p \leq K} \mu_p(E_{p,v_p(n),v_p(m)}).$$

In other words, by the characterization (5.1) of $E_{p,\alpha,\beta}$,

$$\lim_{T \rightarrow +\infty} \mathbb{P}_T(\forall p \leq K, v_p(u_\gamma) = v_p(n), v_p(f_\gamma) = v_p(m)) = \mu(G_{n,m,K}),$$

so for all T large enough,

$$|\mathbb{P}_T(\forall p \leq K, v_p(u_\gamma) = v_p(n), v_p(f_\gamma) = v_p(m)) - \mu(G_{n,m,K})| \leq \epsilon.$$

But since $u_\gamma f_\gamma$ has no large prime factor $p \geq K$ with probability $1 - \epsilon$, for T large enough

$$\mathbb{P}_T\left(u_\gamma = \prod_{p \leq K} p^{v_p(u_\gamma)}, f_\gamma = \prod_{p \leq K} p^{v_p(f_\gamma)}\right) \geq 1 - \epsilon.$$

Since we chose K to be larger than any prime factor of nm , we have $n = \prod_{p \leq K} p^{v_p(n)}$ and similarly for m . The last two inequalities imply:

$$|\mathbb{P}_T(u_\gamma = n, f_\gamma = m) - \mu(G_{n,m,K})| \leq 2\epsilon.$$

Recall that K was chosen in order to satisfy Inequality (5.2), which can be rewritten:

$$|\mu(G_{n,m,K}) - q_{n,m}| \leq \epsilon,$$

so we get for all T large enough,

$$|\mathbb{P}_T(u_\gamma = n, f_\gamma = m) - q_{n,m}| \leq 3\epsilon,$$

which concludes the proof of the convergence part of Theorem 1.3. Now, we wish to check that $(q_{n,m})_{n,m \geq 1}$ defines a probability measure. Since for fixed p , the sets $(E_{p,\alpha,\beta})_{\alpha,\beta \geq 0}$ partition $\text{SL}(2, \mathbb{Z}_p)$ modulo a set of zero μ_p -measure, we have

$$\sum_{\alpha,\beta \geq 0} \mu_p(E_{p,\alpha,\beta}) = 1.$$

A sequence $(\alpha_p)_p$ of integers indexed by prime numbers is almost-zero (a.z.) if $\alpha_p = 0$ for sufficiently large p . With this notation, because of the convergence of $\prod_p \mu_p(E_{p,0,0})$, it follows that for any sequences $(\alpha_p)_p, (\beta_p)_p \geq 0$, the product $\prod_p \mu_p(E_{p,\alpha_p,\beta_p})$ is nonzero if and only if both sequences $(\alpha_p)_p, (\beta_p)_p$ are almost-zero: indeed, since $\lim_{p \rightarrow \infty} \mu_p(E_{p,0,0}) = 1$, we have $\inf_p \mu_p(E_{p,0,0}) > 0$ so

$$\delta := \sup_{p, (\alpha,\beta) \neq (0,0)} \mu_p(E_{p,\alpha,\beta}) \leq 1 - \inf_p \mu_p(E_{p,0,0}) < 1.$$

It follows that $\prod_p \mu_p(E_{p,\alpha_p,\beta_p}) \leq \delta^k$ when at least k factors correspond to nonzero (α_p, β_p) , so the product must diverge to zero if either $(\alpha_p)_p$ or $(\beta_p)_p$ is not almost zero. This implies the formal identities:

$$1 = \prod_p \left(\sum_{\alpha,\beta \geq 0} \mu_p(E_{p,\alpha,\beta}) \right) = \sum_{(\alpha_p)_p, (\beta_p)_p \text{ a.z.}} \prod_p \mu_p(E_{p,\alpha_p,\beta_p}) = \sum_{n,m \geq 1} q_{n,m}.$$

6. The indexes u_γ , f_γ through reduction mod p^k

6.1. The p -adic valuations of u_γ , $p \neq 2$. Here we express the p -adic valuations of u_γ and f_γ in terms of congruence conditions on the matrix γ , and define the clopen sets $E_{p,\alpha,\beta}$, for an odd prime p . A useful convention is that any congruence condition mod p^0 is automatically satisfied. The case of the prime $p = 2$ is essentially similar, but more complicated so is deferred to Appendix A.

Lemma 6.1 ([7, Lemma 3.1]). *Let $\gamma \in \Gamma$ be hyperbolic, and let p be prime. The p -adic valuation $v_p(u_\gamma)$ of u_γ is the largest integer k such that $\gamma \bmod p^k$ is a scalar matrix. In particular, for $p \neq 2$,*

$$v_p(u_\gamma) = \sup \{k \geq 0 : \gamma = \pm I_2 \bmod p^k\}.$$

Proof. We have

$$\gamma = \begin{pmatrix} a & b \\ c & a + (d - a) \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \bmod u_\gamma,$$

so in particular $\gamma \bmod p^{v_p(u_\gamma)}$ is scalar. Reciprocally, if $\gamma \bmod p^k$ is scalar for some $k \geq 1$, this means that p^k divides the three coefficients $b, c, d - a$, whose gcd is u_γ , so $k \leq v_p(u_\gamma)$. Now we remark that if $p \neq 2$ and $k \geq 1$, there are exactly two scalar matrices in $\mathrm{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$, namely $\pm I_2$. \square

We note for later that being scalar mod p^k implies that the trace is $\pm 2 \bmod p^{2k}$:

Lemma 6.2. *Let $k \geq 1$ and p prime. If $\gamma \in \Gamma$ is such that*

$$\gamma = \pm I_2 \bmod p^k,$$

then

$$\mathrm{Tr}(\gamma) = \pm 2 \bmod p^{2k}.$$

Proof. We may write γ as

$$(6.1) \quad \gamma = \epsilon \left(I_2 + p^k \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right),$$

where a, b, c, d are integers, $\epsilon = \pm 1$. Since γ has determinant 1, we have

$$(a + d)p^k + (ad - bc)p^{2k} = 0.$$

We factor p^k , and reduce modulo p^k :

$$a + d = 0 \bmod p^k,$$

Now returning to Equation (6.1) and taking the trace, we get

$$\begin{aligned} \mathrm{Tr}(\gamma) &= \epsilon(2 + (a + d)p^k) \\ &= \pm 2 \bmod p^{2k}. \end{aligned}$$

\square

6.2. The p -adic valuations of $u_\gamma f_\gamma$, $p \neq 2$.

Lemma 6.3. *If $p \neq 2$ is prime, then*

$$v_p(u_\gamma f_\gamma) = \sup\{k \geq 0 : \text{Tr}(\gamma) = \pm 2 \bmod p^{2k}\}.$$

Proof. Let D be the square-free positive integer such that $\mathbb{K} = \mathbb{Q}(\sqrt{D})$, and recall that

$$\Delta_{\mathbb{K}} = \begin{cases} 4D & \text{if } D \equiv 2, 3 \pmod{4}, \\ D & \text{if } D \equiv 1 \pmod{4}. \end{cases}$$

Since $\text{Tr}(\gamma)^2 - 4 = (u_\gamma f_\gamma)^2 \Delta_{\mathbb{K}}$, and $p \neq 2$,

$$v_p(\text{Tr}(\gamma)^2 - 4) = \begin{cases} 1 + 2v_p(u_\gamma f_\gamma) & \text{if } D \equiv 0 \pmod{p}, \\ 2v_p(u_\gamma f_\gamma) & \text{if } D \not\equiv 0 \pmod{p}. \end{cases}$$

But $\text{Tr}(\gamma)^2 - 4 = (\text{Tr}(\gamma) - 2)(\text{Tr}(\gamma) + 2)$ where $\text{Tr}(\gamma) + 2$ and $\text{Tr}(\gamma) - 2$ have a gcd dividing 4, so at least one of the two terms $\text{Tr}(\gamma) \pm 2$ is not divisible by $p \neq 2$. It follows that:

$$v_p(\text{Tr}(\gamma)^2 - 4) = \max(v_p(\text{Tr}(\gamma) - 2), v_p(\text{Tr}(\gamma) + 2)),$$

so $v_p(u_\gamma f_\gamma)$ is the largest integer k such that one of the two numbers $\text{Tr}(\gamma) \pm 2$ is divisible by p^{2k} . \square

We now define for $p \neq 2$ and $\alpha, \beta \geq 0$ the set

$$F_{p,\alpha,\beta} = \left\{ \gamma \in \text{SL}(2, \mathbb{Z}_p) : \gamma \equiv \pm I_2 \pmod{p^\alpha}, \text{Tr}(\gamma) \equiv \pm 2 \pmod{p^{2\beta}} \right\}.$$

With the convention on congruences mod p^0 , we have $F_{p,0,0} = \text{SL}(2, \mathbb{Z}_p)$. Clearly, this set is defined mod $p^{\max(\alpha, 2\beta)}$ so is clopen in $\text{SL}(2, \mathbb{Z}_p)$. These sets are decreasing in both variables: if $\alpha' \geq \alpha, \beta' \geq \beta$, then $F_{p,\alpha',\beta'} \subset F_{p,\alpha,\beta}$. By the Lemmata 6.1 and 6.3, for any hyperbolic matrix $\gamma \in \Gamma$,

$$v_p(u_\gamma) \geq \alpha, v_p(u_\gamma f_\gamma) \geq \beta \iff \gamma \in F_{p,\alpha,\beta}.$$

Define the clopen set

$$H_{p,\alpha,\beta} = F_{p,\alpha,\beta} - F_{p,\alpha+1,\beta},$$

so we have for any hyperbolic matrix $\gamma \in \Gamma$

$$v_p(u_\gamma) = \alpha, v_p(f_\gamma) \geq \beta - \alpha \iff \gamma \in H_{p,\alpha,\beta}.$$

Notice that $H_{p,\alpha,\beta'} \subset H_{p,\alpha,\beta}$ if $\beta' \geq \beta$.

Now for $i \geq 0, j \geq 0$, we define the (still clopen) sets

$$E_{p,\alpha,\beta} = H_{p,\alpha,\alpha+\beta} - H_{p,\alpha,\alpha+\beta+1},$$

so this time

$$v_p(u_\gamma) = \alpha, v_p(f_\gamma) = \beta \iff \gamma \in E_{p,\alpha,\beta}.$$

The sets $(E_{p,\alpha,\beta})_{\alpha,\beta \geq 0}$ are disjoint: in fact we could have defined them directly as the set of matrices $\gamma \in \text{SL}(2, \mathbb{Z}_p)$ such that γ is congruent to $\pm I_2$

mod p^α but not $p^{\alpha+1}$, and $\text{Tr}(\gamma)$ is congruent to $\pm 2 \bmod p^{2(\alpha+\beta)}$ but not mod $p^{2(\alpha+\beta)+2}$. But it will be more convenient to decompose them like this for counting purposes. Matrices $\gamma \in \text{SL}(2, \mathbb{Z}_p)$ that do not belong to any $E_{p,\alpha,\beta}$ are either equal to $\pm I_2$, or such that their trace is ± 2 . Since the first possibility also imply the second, we have:

$$\text{SL}(2, \mathbb{Z}_p) - \text{Tr}^{-1}(\pm 2) = \bigsqcup_{\alpha, \beta \geq 0} E_{p,\alpha,\beta},$$

as claimed.

7. Counting matrices in $\text{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$

In this section, we count in an elementary fashion the number of matrices in $\text{SL}(2, \mathbb{Z}/p^k\mathbb{Z})$ satisfying some congruence condition, for an odd prime p . The objectives are:

- To obtain formulas for the p -adic measures $\mu_p(E_{p,\alpha,\beta})$ that are involved in the definition of $q_{n,m}$. The first step will be to compute $\mu_p(E_{p,\alpha,\beta})$ for $p \neq 2$.
- To check that the set of matrices of trace ± 2 in $\text{SL}(2, \mathbb{Z}_p)$ is of zero μ_p measure, which is the last unproven claim of the first fact of Section 5.1. This is a corollary of the previous step, as one can check using the formulas obtained that

$$\sum_{\alpha, \beta \geq 0} \mu_p(E_{p,\alpha,\beta}) = 1.$$

- To check the second claimed fact of Section 5.1, that is the convergence of the infinite product

$$q_{1,1} = \prod_p \mu_p(E_{p,0,0}).$$

- To obtain the formulas for $\sum_{m \geq 1} q_{1,m}$, $\sum_{n \geq 1} q_{n,1}$ appearing in Corollaries 1.4 and 1.5.

Again, the bothersome case of the prime $p = 2$ is treated in the appendix.

7.1. Preliminaries. We recall the following:

Lemma 7.1 ([12, Theorem 4.2.4]). *Let p be a prime number, and $k \geq 1$. Then*

$$|\text{SL}(2, \mathbb{Z}/p^k\mathbb{Z})| = p^{3k-2}(p^2 - 1).$$

We will also need the following:

Lemma 7.2. *Let p be a prime number and $k \geq 1$. The number of points $(X, Y, Z) \in (\mathbb{Z}/p^k\mathbb{Z})^3$ on the quadric surface defined by*

$$X^2 + YZ = 0,$$

is

$$N_{p,k} = p^{2k} + p^{2k-1} - p^{\lfloor \frac{3k-1}{2} \rfloor}.$$

Proof. We count the solutions in an elementary fashion. Given $m \geq 0$, denote by $N_{p,k,m}$ the number of solutions such that $v_p(Y) = m$. Notice first that in the case where $Y = 0$, then $X^2 = 0$ necessarily so $v_p(X) \geq \lceil k/2 \rceil$. Thus if $r = \lceil k/2 \rceil$, X can be written $X = p^r U$ with $U \in \mathbb{Z}/p^{k-\lceil k/2 \rceil}\mathbb{Z}$ arbitrary, and Z can be chosen arbitrarily. Thus

$$N_{p,k,\infty} = p^{2k-\lceil k/2 \rceil}.$$

We now assume that $0 \leq m \leq k-1$ i.e. $Y \neq 0$.

If (X, Y, Z) is a solution, then from $X^2 = -YZ$ we get that $2v_p(X) \geq v_p(Y) = m$. Let $r = \lceil m/2 \rceil$ be the least integer larger than $m/2$, thus if (X, Y, Z) is a solution with $v_p(Y) = m$, then there exists $U \in \mathbb{Z}/p^{k-r}\mathbb{Z}$, $V \in (\mathbb{Z}/p^{k-m}\mathbb{Z})^*$ such that

$$X = p^r U, \quad Y = p^m V$$

We can rewrite the equation $X^2 + YZ = 0$ as

$$p^{2r} U^2 = -p^m V Z \pmod{p^k},$$

equivalently

$$Z = -p^{2r-m} U^2 V^{-1} \pmod{p^{k-m}},$$

since V has an inverse mod p^{k-m} . Thus given any pair $(U, V) \in (\mathbb{Z}/p^{k-r}\mathbb{Z}) \times (\mathbb{Z}/p^{k-m}\mathbb{Z})^*$, there are exactly p^m values of $Z \in \mathbb{Z}/p^k\mathbb{Z}$ such that $(X = p^r U, Y = p^m V, Z)$ is a solution to $X^2 + YZ = 0$. We have obtained that

$$N_{p,k,m} = p^{k-r} (p^{k-m} - p^{k-m-1}) p^m = p^{2k-\lceil m/2 \rceil} - p^{2k-\lceil m/2 \rceil-1}$$

Summing over m gives

$$N_{p,k} = N_{p,k,\infty} + \sum_{m=0}^{k-1} N_{p,k,m},$$

Notice that the subsequence for even terms is telescopic, so we get

$$\sum_{m=2n \leq k-1} N_{p,k,m} = p^{2k} - p^{2k-\lfloor (k-1)/2 \rfloor-1},$$

because if $m = 2n$ is even, then $\lceil m/2 \rceil = n$ and the sum runs from $n = 0$ to $n = \lfloor (k-1)/2 \rfloor$. Similarly for odd terms

$$\sum_{m=2n-1 \leq k-1} N_{p,k,m} = p^{2k-1} - p^{2k-\lfloor k/2 \rfloor-1},$$

since if $m = 2n - 1$ is odd, then $\lceil m/2 \rceil = n$, and the sum runs from $n = 1$ to $n = \lfloor k/2 \rfloor$. Note also that in the case where the sum is over an empty set, i.e. $\lfloor k/2 \rfloor < 1$, then $k = 1$ and the above formula is still valid in this case. Thus,

$$N_{p,k} = p^{2k} + p^{2k-1} + p^{2k-\lceil k/2 \rceil} - p^{2k-\lfloor (k-1)/2 \rfloor - 1} - p^{2k-\lfloor k/2 \rfloor - 1}.$$

Two of the last three terms cancel out according to the parity of k . In both cases, we can write;

$$N_{p,k} = p^{2k} + p^{2k-1} - p^{\lfloor \frac{3k-1}{2} \rfloor}.$$

□

Lemma 7.3. *Let $\alpha \geq 0, \beta \geq 0$ with $(\alpha, \beta) \neq (0, 0)$, and p an odd prime. The number of matrices γ in $\mathrm{SL}(2, \mathbb{Z}/p^{2(\alpha+\beta)}\mathbb{Z})$ such that*

$$\begin{cases} \gamma = \pm I_2 \bmod p^\alpha, \\ \mathrm{Tr}(\gamma) = \pm 2 \bmod p^{2(\alpha+\beta)} \end{cases}$$

is

$$2p^{3\alpha}(p^{4\beta} + p^{4\beta-1} - p^{3\beta-1}),$$

Proof. First of all, since $\mathrm{Tr}(I_2) = +2$, if $\alpha \geq 1$, the sign must be the same in the two defining congruences. If $\alpha = 0$ then the first congruence does not matter. So it is enough to count the matrices such that

$$\begin{cases} \gamma = +I_2 \bmod p^\alpha, \\ \mathrm{Tr}(\gamma) = +2 \bmod p^{2(\alpha+\beta)} \end{cases}$$

and multiply the result by 2, as $-\gamma$ satisfy opposite congruences, and $-2 \not\equiv +2 \bmod p^{2(\alpha+\beta)}$ since p is odd and $\alpha + \beta \geq 1$.

We can write any matrix in $\mathrm{SL}(2, \mathbb{Z}/p^{2(\alpha+\beta)}\mathbb{Z})$ satisfying the above congruences as

$$\gamma = \begin{pmatrix} 1 + p^\alpha a & p^\alpha b \\ p^\alpha c & 1 - p^\alpha a \end{pmatrix},$$

where $(a, b, c) \in \mathbb{Z}/p^{\alpha+2\beta}\mathbb{Z}$. Given such (a, b, c) , the matrix above is in $\mathrm{SL}(2, \mathbb{Z}/p^{2(\alpha+\beta)}\mathbb{Z})$ iff

$$(1 - p^\alpha a)(1 + p^\alpha a) - p^{2\alpha}bc = 1 \bmod p^{2(\alpha+\beta)},$$

that is

$$(7.1) \quad a^2 + bc = 0 \bmod p^{2\beta}.$$

By Lemma 7.2, this quadric has $p^{4\beta} + p^{4\beta-1} - p^{3\beta-1}$ points in $\mathbb{Z}/p^{2\beta}\mathbb{Z}$. Each element of $\mathbb{Z}/p^{2\beta}\mathbb{Z}$ has p^α lifts in $\mathbb{Z}/p^{\alpha+2\beta}\mathbb{Z}$, so there are $p^{3\alpha}(p^{4\beta} + p^{4\beta-1} - p^{3\beta-1})$ triplets (a, b, c) in $\mathbb{Z}/p^{\alpha+2\beta}\mathbb{Z}$ satisfying (7.1). □

7.2. Measure of $E_{p,\alpha,\beta}$, $p \neq 2$. Recall that we defined for odd primes p :

$$F_{p,\alpha,\beta} = \left\{ \gamma \in \mathrm{SL}(2, \mathbb{Z}_p) : \gamma = \pm I_2 \bmod p^\alpha, \mathrm{Tr}(\gamma) = \pm 2 \bmod p^{2\beta} \right\}.$$

Lemma 7.4. *Let $p \neq 2$ be prime, and $\alpha \geq 0$, $\beta \geq 0$ be integers,*

$$\mu_p(F_{p,\alpha,\beta}) = \begin{cases} 1 & \text{if } (\alpha, \beta) = (0, 0), \\ \frac{2}{p^{3\alpha-2}(p^2-1)} & \text{if } (\alpha, \beta) \neq (0, 0), \beta \leq \alpha, \\ \frac{2(p^{\beta-\alpha+1} + p^{\beta-\alpha} - 1)}{p^{3\beta-1}(p^2-1)} & \text{if } (\alpha, \beta) \neq (0, 0), \beta \geq \alpha. \end{cases}$$

Proof. For $(\alpha, \beta) = (0, 0)$, by convention, $F_{p,0,0} = \mathrm{SL}(2, \mathbb{Z}_p)$ so

$$\mu_p(F_{p,0,0}) = 1.$$

If $\beta \leq \alpha$, by Lemma 6.2, we have $F_{p,\alpha,\beta} = F_{p,\alpha,\alpha}$; note that the two applicable formulas given in the statement agree when $\beta = \alpha$, so it is sufficient to prove the formula when $\beta \geq \alpha$.

In this case,

$$\mu_p(F_{p,\alpha,\beta}) = \frac{\left| \left\{ \gamma \in \mathrm{SL}(2, \mathbb{Z}/p^{2\beta}\mathbb{Z}) : \gamma = \pm I_2 \bmod p^\alpha, \mathrm{Tr}(\gamma) = \pm 2 \bmod p^{2\beta} \right\} \right|}{|\mathrm{SL}(2, \mathbb{Z}/p^{2\beta}\mathbb{Z})|},$$

and by Lemmata 7.1 and 7.3, this ratio is equal to

$$\mu_p(F_{p,\alpha,\beta}) = \frac{2p^{3\alpha}(p^{4(\beta-\alpha)} + p^{4(\beta-\alpha)-1} - p^{3(\beta-\alpha)-1})}{p^{3\beta-2}(p^2-1)}. \quad \square$$

Proposition 7.5. *For $p \neq 2$, $\alpha \geq 0$, $\beta \geq 0$,*

$$\mu_p(E_{p,\alpha,\beta}) = \begin{cases} 1 - \frac{2(p^2 + p - 1)}{p^2(p^2 - 1)} & \text{if } (\alpha, \beta) = (0, 0), \\ \frac{2(p^3 + p^2 - 1)}{p^{3\alpha+2}(p+1)} & \text{if } \alpha > 0, \beta = 0, \\ \frac{2(p^2 - 1)}{p^{3\alpha+2\beta+2}} & \text{if } \beta > 0. \end{cases}$$

Proof. Recall that $H_{p,\alpha,\beta}$ was defined as $F_{p,\alpha,\beta} - F_{p,\alpha+1,\beta}$, and that $F_{p,\alpha+1,\beta} \subset F_{p,\alpha,\beta}$, so

$$\mu_p(H_{p,\alpha,\beta}) = \mu_p(F_{p,\alpha,\beta}) - \mu_p(F_{p,\alpha+1,\beta}).$$

Likewise, $E_{p,\alpha,\beta} = H_{p,\alpha,\alpha+\beta} - H_{p,\alpha,\alpha+\beta+1}$ where $H_{p,\alpha,\alpha+\beta+1} \subset H_{p,\alpha,\alpha+\beta}$. So we get the inclusion-exclusion-like formula:

$$\begin{aligned} \mu_p(E_{p,\alpha,\beta}) &= \mu_p(F_{p,\alpha,\alpha+\beta}) + \mu_p(F_{p,\alpha+1,\alpha+\beta+1}) \\ &\quad - \mu_p(F_{p,\alpha+1,\alpha+\beta}) - \mu_p(F_{p,\alpha,\alpha+\beta+1}). \end{aligned}$$

We first consider the case where $\beta > 0$. Then $\alpha + \beta \geq \alpha + 1$ so we can consistently apply the third formula for the terms $\mu_p(F_{p,\alpha',\beta'})$:

$$\begin{aligned}\mu_p(E_{p,\alpha,\beta}) &= 2 \left(\frac{p^3(p^{\beta+1} + p^\beta - 1) + (p^{\beta+1} + p^\beta - 1)}{p^{3\alpha+3\beta+2}(p^2 - 1)} \right. \\ &\quad \left. - \frac{p^3(p^\beta + p^{\beta-1} - 1) + (p^{\beta+2} + p^{\beta+1} - 1)}{p^{3\alpha+3\beta+2}(p^2 - 1)} \right), \\ &= \frac{2(p^{\beta+4} - 2p^{\beta+2} + p^\beta)}{p^{3\alpha+3\beta+2}(p^2 - 1)} \\ &= \frac{2(p^2 - 1)}{p^{3\alpha+2\beta+2}}.\end{aligned}$$

We now consider the subcase where $\beta = 0$. In this case, the two terms $\mu_p(F_{p,\alpha+1,\alpha+1})$ and $\mu_p(F_{p,\alpha+1,\alpha})$ cancel each other so the formula reduces to

$$\mu_p(E_{p,\alpha,0}) = \mu_p(F_{p,\alpha,\alpha}) - \mu_p(F_{p,\alpha,\alpha+1}).$$

If $\alpha = 0$ also, we get

$$\mu_p(E_{p,0,0}) = 1 - \mu_p(F_{p,0,1}) = 1 - \frac{2(p^2 + p - 1)}{p^2(p^2 - 1)}.$$

If $\alpha > 0$ and $\beta = 0$ then

$$\begin{aligned}\mu_p(E_{p,\alpha,0}) &= \frac{2}{p^{3\alpha-2}(p^2 - 1)} - \frac{2(p^2 + p - 1)}{p^{3\alpha+2}(p^2 - 1)}, \\ &= \frac{2(p^4 - p^2 - p + 1)}{p^{3\alpha+2}(p^2 - 1)}, \\ &= \frac{2(p^3 + p^2 - 1)}{p^{3\alpha+2}(p + 1)}.\end{aligned}\quad \square$$

7.3. Corollaries. Since $\mu_p(E_{p,0,0}) = 1 + O(p^{-2})$, we have obtained the awaited second fact:

Corollary 7.6. *The product $\prod_p \mu_p(E_{p,0,0})$ converges.*

We leave to the reader to check that $\sum_{\alpha,\beta} \mu_p(E_{p,\alpha,\beta}) = 1$, which implies that these sets cover $\mathrm{SL}(2, \mathbb{Z}_p)$ up to a set of measure zero, as claimed. Let's find the infinite products appearing in Corollaries 1.4 and 1.5.

Corollary 7.7.

$$\sum_{n \geq 1} q_{1,m} = \frac{5}{6} \prod_{p \geq 3} \left(1 - \frac{2}{p(p^2 - 1)} \right).$$

Proof. We have

$$\sum_{m \geq 1} q_{1,m} = \sum_m \prod_p \mu_p(E_{p,0,v_p(m)}) = \prod_p \left(\sum_{\beta \geq 0} \mu_p(E_{p,0,\beta}) \right),$$

so we can compute first the sum for a fixed p . For $p = 2$, $\bigsqcup_{\beta \geq 0} E_{2,0,\beta} = A_0$ is the set of matrices not congruent to $I_2 \bmod 2$, has measure $5/6$ because out of the 6 matrices in $\mathrm{SL}(2, \mathbb{Z}/2\mathbb{Z})$, only one is scalar (this can also be obtained but summing the expressions given in Proposition A.4). For $p > 2$,

$$\sum_{\beta \geq 0} \mu_p(E_{p,0,\beta}) = \mu_p(F_{p,0,0}) - \mu_p(F_{p,0,1}),$$

quantities for which Lemma 7.4 gives formulas. \square

Corollary 7.8.

$$\sum_{n \geq 1} q_{n,1} = \frac{75}{112} \prod_{p \geq 3} \left(1 - \frac{2p}{p^3 - 1} \right).$$

Proof. Likewise,

$$\sum_{n \geq 1} q_{n,1} = \prod_p \left(\sum_{\alpha \geq 0} \mu_p(E_{p,\alpha,0}) \right).$$

For the prime 2, from Proposition A.4, we get by summing over α :

$$\mu_2 \left(\bigsqcup_{\alpha \geq 0} E_{2,\alpha,0} \right) = \frac{75}{112}.$$

For primes $p \geq 3$, we get

$$\sum_{\alpha \geq 0} \mu_p(E_{p,\alpha,0}) = 1 - \frac{2(p^2 + p - 1)}{p^2(p^2 - 1)} + \frac{2(p^3 + p^2 - 1)}{p^2(p + 1)(p^3 - 1)} = 1 - \frac{2p}{p^3 - 1}. \quad \square$$

8. Large prime factors of $u_\gamma f_\gamma$

In this section, we prove the third ingredient of Theorem 1.3:

Proposition 8.1. *For any $\epsilon > 0$, there exists $K(\epsilon) > 0$ such that the upper density of the set*

$$M_{K(\epsilon)} = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}) : |\mathrm{Tr}(\gamma)| > 2, \exists p \geq K(\epsilon) \text{ prime, } p|u_\gamma \text{ or } p|f_\gamma\},$$

is smaller than ϵ .

Proof. We first fix $K \geq 3$ and compute an upper bound for the upper density of M_K ; it will be sufficient to check that this bound does tend to zero as $K \rightarrow +\infty$.

If $p > 2$ is prime, $\gamma \in \Gamma$ hyperbolic, and p divides either u_γ or f_γ , then since

$$\mathrm{Tr}(\gamma)^2 - 4 = u_\gamma^2 f_\gamma^2 \Delta_{\mathbb{K}},$$

we see that p^2 divides the product $\text{Tr}(\gamma)^2 - 4 = (t - 2)(t + 2)$, where $t := \text{Tr}(\gamma)$. This means that

$$t = \pm 2 \bmod p^2,$$

because the gcd of $t - 2$ and $t + 2$ divides 4. This congruence of the trace merely depends on the (right) coset of $\gamma \bmod \Gamma(p^2)$. Denote by $Z_p \subset \Gamma$ a set of representatives of cosets $\bmod \Gamma(p^2)$ whose trace are $\pm 2 \bmod p^2$, so that

$$M_K \subset \bigcup_{p \geq K} Z_p \Gamma(p^2).$$

Let $T > K$ be large, and $\alpha \in (0, 1)$ be a (small) parameter to be chosen later. We cut $M_K \cap \Gamma_T$ in two parts, the part M_K^{small} where $t^2 - 4$ is divisible by p^2 for some relatively small $p \leq T^\alpha$ (but still $p \geq K$), and the part M_K^{big} where $t^2 - 4$ has a very large square prime factors $p \geq T^\alpha$.

We first analyze M_K^{small} . Let p be a small prime $p \leq T^\alpha$, $p \geq K$. By the previous remark,

$$M_K^{small} \cap \Gamma_T \subset \bigcup_{K \leq p \leq T^\alpha} (Z_p \Gamma(p^2) \cap \Gamma_T).$$

By the uniform equidistribution Theorem 4.1 ([15]), there exist constants $C > 0$, $\beta > 0$, $T_0 > 0$ such that for all p and $\gamma_0 \in \Gamma$ and $T \geq T_0$,

$$\left| |\gamma_0 \Gamma(p^2) \cap \Gamma_T| - \frac{|\Gamma_T|}{[\Gamma : \Gamma(p^2)]} \right| \leq C |\Gamma_T|^{1-\beta},$$

so

$$|\gamma_0 \Gamma(p^2) \cap \Gamma_T| \leq \frac{|\Gamma_T|}{[\Gamma : \Gamma(p^2)]} + C |\Gamma_T|^{1-\beta}.$$

Thus, adding for each $p \in (K, T^\alpha)$ and each $\gamma_0 \in Z_p$ these upper bounds, we get

$$|M_K^{small}| \leq \sum_{K \leq p \leq T^\alpha} |Z_p| \left(\frac{|\Gamma_T|}{[\Gamma : \Gamma(p^2)]} + C |\Gamma_T|^{1-\beta} \right).$$

By Lemma 7.3 applied with $\alpha = 0$, $\beta = 1$, we get the estimate

$$|Z_p| = 2(p^4 + p^3 - p^2) \leq 4p^4,$$

and by Lemma 7.1, $[\Gamma : \Gamma(p^2)] = p^4(p^2 - 1) \geq p^6/2$. Thus

$$|M_K^{small}| \leq |\Gamma_T| \left(\sum_{K \leq p \leq T^\alpha} \frac{8}{p^2} \right) + C \left(\sum_{K \leq p \leq T^\alpha} 4p^4 \right) |\Gamma_T|^{1-\beta}.$$

We can bound $\sum_{K \leq p \leq T^\alpha} 4p^4 \leq 4T^{5\alpha}$. Also, since $|\Gamma_T| \sim 6T^2$, for T large enough, $|\Gamma_T| \geq 5T^2$, so we have

$$\mathbb{P}_T(M_K^{small}) \leq \frac{8}{K-1} + \frac{C}{5} T^{5\alpha-2\beta}.$$

We now fix the choice of α , such that $5\alpha < 2\beta$. We get the desired conclusion that M_K^{small} has small \mathbb{P}_T measure when K and T are large.

Now we consider M_K^{big} . For fixed p , because $|t| \leq 2T$, there are at most $8T/p^2$ possible traces t that satisfies the congruence condition $t \equiv \pm 2 \pmod{p^2}$ (recall that we excluded $t \equiv \pm 2$). By Lemma 3.1, each such trace is realized by at most $c_1 T^{1+\eta}$ matrices in Γ_T , where $\eta > 0$ is chosen such that $\eta < \alpha$. To summarize,

$$|M_K^{big}| \leq \sum_{p \geq T^\alpha} \frac{8T}{p^2} c_1 T^{1+\eta} \leq c_2 T^{2+\eta} \sum_{p \geq T^\alpha} \frac{1}{p^2}.$$

One has $\sum_{p \geq T^\alpha} \frac{1}{p^2} = O(T^{-\alpha})$, so

$$|M_K^{big}| \leq c_3 T^{2+\eta-\alpha}.$$

We thus obtain

$$\mathbb{P}_T(M_K^{big}) = O(T^{\eta-\alpha}). \quad \square$$

9. Class number

In this section, we indicate how to prove Corollary 1.7, which states that class numbers are of the order of $T^{1 \pm \epsilon}$ with high probability.

By the Brauer-Siegel Theorem, when we consider sequences of quadratic fields \mathbb{K} ,

$$\frac{\log(h(\Delta_{\mathbb{K}}) \log(\epsilon_{\mathbb{K}}))}{\log \sqrt{\Delta_{\mathbb{K}}}} \rightarrow 1,$$

as the fundamental discriminant $\Delta_{\mathbb{K}}$ of the quadratic field tends to infinity. Now pick a random matrix γ using \mathbb{P}_T , which is likely hyperbolic so we can consider the quadratic field $\mathbb{K} = \mathbb{Q}[\sqrt{\Delta_\gamma}] = \mathbb{Q}[\epsilon_\gamma]$. By Theorem 1.2 and Proposition 1.1, it is likely that $\epsilon_{\mathbb{K}} = |\epsilon_\gamma|$ is of the order of T in the sense that for every $\eta > 0$, there exists a $c > 0$ such that on a set of measure $1 - \eta$,

$$cT < \epsilon_{\mathbb{K}} < T.$$

By Theorem 2, $\Delta_{\mathbb{K}}$ is of the order of T^2 with high probability, in the same sense. Embedding these two estimates in the Brauer-Siegel Theorem, we obtain that with high probability, $h(\Delta_{\mathbb{K}})$ is generally of size $T^{1 \pm \epsilon}$: for all $\epsilon > 0$,

$$\lim_{T \rightarrow \infty} \mathbb{P}_T \left(\left| \frac{\log(h(\Delta_{\mathbb{K}}))}{\log T} - 1 \right| < \epsilon \right) = 1.$$

There is a classical formula for $h(\Delta_\gamma)$, see for example [14, Theorem 12.12], or Conrad's notes [4]:

$$\frac{h(\mathcal{O}_\gamma)}{h(\mathcal{O}_{\mathbb{K}})} = \frac{[\mathcal{O}_{\mathbb{K}}/f_\gamma \mathcal{O}_{\mathbb{K}} : \mathcal{O}_\gamma/f_\gamma \mathcal{O}_{\mathbb{K}}]}{[\mathcal{O}_{\mathbb{K}}^\times : \mathcal{O}_\gamma^\times]}.$$

Here $h(\mathcal{O})$ denotes is the classical (not narrow) class number of the ring \mathcal{O} , that may be equal or half the corresponding narrow class number $h(\Delta)$. Anyway, the formula shows that the ratio $h(\Delta_\gamma)/h(\Delta_\mathbb{K})$ is bounded by $4f_\gamma^2$, so is bounded provided that the conductor f_γ is. Still by Theorem 1.3, f_γ is bounded with large probability, so we also get also in this case: for all $\epsilon > 0$,

$$\lim_{T \rightarrow \infty} \mathbb{P}_T \left(\left| \frac{\log(h(\Delta_\gamma))}{\log T} - 1 \right| < \epsilon \right) = 1.$$

Appendix A. The prime $p = 2$

In this subsection, we consider the prime $p = 2$. The first task is to identify the 2-adic valuation of u_γ and f_γ by congruence conditions, in order to identify the sets $E_{2,\alpha,\beta}$. The second task is to compute their measure. These two tasks involve some tedious case-by-case analysis.

A.1. The 2-adic valuation of u_γ . By Lemma 6.1, the 2-adic valuation of u_γ is the largest power 2^k such that $\gamma \bmod 2^k$ is scalar. However, scalar matrices in this case are not limited to $\pm I_2$:

Lemma A.1. *If $p = 2$, the scalar matrices in $\mathrm{SL}(2, \mathbb{Z}/2^k\mathbb{Z})$ are*

- I_2 if $k = 1$,
- $\pm I_2$ if $k = 2$,
- $\pm I_2$ and $\pm(1 + 2^{k-1})I_2$ if $k \geq 3$.

The proof is left to the reader. It will be useful for later to note that being congruent to either $\pm I_2$ or one of these two additional scalar matrices mod 2^k can be distinguished by looking at the trace mod 2^{2k} .

Lemma A.2. *Let $k \geq 3$. If*

$$\gamma = \pm(1 + 2^{k-1})I_2 \bmod 2^k,$$

then

$$\mathrm{Tr}(\gamma) = \pm(2 + 2^{2k-2}) \bmod 2^{2k}.$$

Proof. This is pretty similar to the proof of Lemma 6.2. We may write γ as

$$(A.1) \quad \gamma = \epsilon \left((1 + 2^{k-1})I_2 + 2^k \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right),$$

where a, b, c, d are integers, and $\epsilon = \pm 1$. Since γ has determinant 1, we have

$$(1 + 2^{k-1} + a2^k)(1 + 2^{k-1} + d2^k) - 2^{2k}bc = 1.$$

So

$$2^k + 2^{2k-2} + (a + d)(1 + 2^{k-1})2^k + (ad - bc)2^{2k} = 0.$$

We factor 2^k , then reduce modulo 2^k :

$$(1 + 2^{k-2}) + (a + d)(1 + 2^{k-1}) = 0 \bmod 2^k,$$

whose solution for $k \geq 3$ is:

$$a + d = -1 + 2^{k-2} \bmod 2^k.$$

Now returning to Equation (A.1) and taking the trace, we get

$$\begin{aligned} \text{Tr}(\gamma) &= \epsilon(2 + (a + d + 1)2^k) \\ &= \pm(2 + 2^{2k-2}) \bmod 2^{2k}. \end{aligned} \quad \square$$

A.2. The 2-adic valuation of $u_\gamma f_\gamma$. We now turn our attention to the 2-adic valuation of $u_\gamma f_\gamma$, which can be determined by looking at the trace of γ as follows:

Lemma A.3. *Let $p = 2$, and $t = \text{Tr}(\gamma)$ where $\gamma \in \Gamma$ is hyperbolic. Then*

- $v_2(u_\gamma f_\gamma) = 0$ if and only if $t \bmod 16 \notin \{2, 6, 10, 14\}$.
- $v_2(u_\gamma f_\gamma) = 1$ if and only if $t \bmod 16 \in \{6, 10\}$.
- $v_2(u_\gamma f_\gamma) = 2$ if and only if $t \bmod 64 \in \{14, 30, 34, 50\}$.
- For $\alpha \geq 3$, $v_2(u_\gamma f_\gamma) = \alpha$ if and only if t is of the form

$$\begin{aligned} &\pm(2 + 2^{2\alpha+1}) \bmod 2^{2\alpha+2}, \\ &\pm(2 + 3 \cdot 2^{2\alpha}) \bmod 2^{2\alpha+2}, \\ &\text{or } \pm(2 + 2^{2\alpha-2}) \bmod 2^{2\alpha}. \end{aligned}$$

Proof. Let D be the square-free positive integer such that $\mathbb{K} = \mathbb{Q}(\sqrt{D})$, and recall that

$$\Delta_{\mathbb{K}} = \begin{cases} 4D & \text{if } D \equiv 2, 3 \pmod{4}, \\ D & \text{if } D \equiv 1 \pmod{4}. \end{cases}$$

Since $\text{Tr}(\gamma)^2 - 4 = (u_\gamma f_\gamma)^2 \Delta_{\mathbb{K}}$, we have,

$$(A.2) \quad v_2(t^2 - 4) = \begin{cases} 2v_2(u_\gamma f_\gamma) & \text{if } D \equiv 1 \pmod{4}, \\ 2v_2(u_\gamma f_\gamma) + 3 & \text{if } D \equiv 2 \pmod{4}, \\ 2v_2(u_\gamma f_\gamma) + 2 & \text{if } D \equiv 3 \pmod{4}. \end{cases}$$

It will be useful to note the following facts:

- The case $D \equiv 2 \pmod{4}$ appears if and only if the valuation of $t^2 - 4$ is odd.
- If the valuation of $t^2 - 4$ is even, i.e. $D \not\equiv 2 \pmod{4}$, then if we factor out the powers of 2, i.e. $t^2 - 4 = 2^{v_p(t^2-4)}w$, then $w \equiv D \pmod{4}$, because 1 is the only odd square mod 4.

We first consider $t \bmod 16$. The following array summarizes what can be said using Equation (A.2), depending on $t \bmod 16$:

| | | | | |
|--------------------------|-----|---------|---------------|----------|
| $t \bmod 16$ | odd | ± 6 | $0, \pm 4, 8$ | ± 2 |
| $D \bmod 4$ | 1 | 2 | 3 | ? |
| $v_p(u_\gamma f_\gamma)$ | 0 | 1 | 0 | ≥ 2 |

Let us show how to find this array. If $t \bmod 16$ is odd, then $t^2 - 4$ is also odd so we must be in the case $D = 1 \bmod 4$ since in all other cases $t^2 - 4$ has valuation ≥ 2 , by (A.2). Hence $v_p(u_\gamma f_\gamma) = 0$.

If $t \bmod 16$ is equal to ± 6 , then $t + 2$ and $t - 2$ are congruent to 8 and 4, or 12 and 8 depending on the sign; in both cases, one of them has 2-adic valuation 3 and the other has valuation 2, so in the end

$$v_2(t^2 - 4) = 5.$$

In the formula (A.2), $D = 2 \bmod 4$ is the only case where the valuation of $t^2 - 4$ is odd, so here we have $2v_2(u_\gamma f_\gamma) + 3 = 5$, so $v_2(u_\gamma f_\gamma) = 1$.

If $t \bmod 16$ is 0, 4, 8 or 12, then $t - 2$ and $t + 2$ both have 2-adic valuation equal to 1, so $v_2(t^2 - 4) = 2$ is even, so D must be 1 or 3 mod 4. Let's show that in fact D must be 3 mod 4. Since $t \bmod 16 \in \{0, 4, 8, 12\}$ so $t^2 - 4 = 12 \bmod 16$ so there exists an integer w such that

$$(A.3) \quad t^2 - 4 = 4w \text{ with } w = 3 \bmod 4.$$

But as noted before, we must have $w = D \bmod 4$, so that $D = 3 \bmod 4$, so by (A.2), $v_2(u_\gamma f_\gamma) = 0$.

We conclude the justification of the array by considering the case where $t = \pm 2 \bmod 16$. In this case, one of the number $t + 2, t - 2$ has valuation at least 4 and the other exactly 2. So the valuation of $t^2 - 4$ is at least 6, and by Equation (A.2),

$$6 \leq v_p(t^2 - 4) \leq 2v_p(u_\gamma f_\gamma) + 3,$$

so $v_p(u_\gamma f_\gamma) \geq 3/2$, but this is an integer so must be ≥ 2 .

The array explains the two first case $v_2(u_\gamma f_\gamma) = 0, 1$ of the statement.

We now consider in more depth the case where $t = \pm 2 \bmod 16$. In this case, since γ is hyperbolic, $t \neq \pm 2$ so there exists a smallest integer $k \geq 2$ such that

$$t \not\equiv \pm 2 \bmod 2^{2k+2}.$$

so there exists an integer $\lambda \in \{1, 2, 3\}$ such that

$$t = \pm(2 + \lambda 2^{2k}) \bmod 2^{2k+2}.$$

Thus for some integer μ ,

$$t = \pm(2 + \lambda 2^{2k} + \mu 2^{2k+2}) \bmod 2^{2k+4},$$

so

$$t^2 = 4 + 4\lambda 2^{2k} \bmod 2^{2k+4},$$

where most of the terms vanished when we expanded the square because $k \geq 2$. Thus,

$$t^2 - 4 = \lambda 2^{2k+2} \bmod 2^{2k+4}.$$

If $\lambda = 2$, then $v_p(t^2 - 4) = 2k + 3$ is odd, so we must be in the case where $D = 2 \bmod 4$, by (A.2), so $v_p(u_\gamma f_\gamma) = k$.

Otherwise, $\lambda = 1, 3$ so $v_p(t^2 - 4) = 2k + 2$ is even so $D = 1, 3 \pmod 4$. Again, $(t^2 - 4)/2^{v_2(t^2 - 4)}$ has the same residue mod 4 as D , so

$$D = \lambda \pmod 4.$$

Still by (A.2), this implies that if $\lambda = 1$, then $v_p(u_\gamma f_\gamma) = k + 1$, and if $\lambda = 3$ then $v_p(u_\gamma f_\gamma) = k$.

Now we can explain the last two lines of the statement. Assume $v_2(u_\gamma f_\gamma) = \alpha \geq 2$. From the above array, t must be $\pm 2 \pmod{16}$.

Define k, λ as above, then $\alpha = k$ if $\lambda = 2, 3$ and $\alpha = k + 1$ if $\lambda = 1$. Since $k \geq 2$, in the case $\alpha = 2$, we cannot have $\lambda = 1$, and so $k = 2$ and thus

$$\alpha = 2 \Rightarrow t = \pm(2 + 16\lambda) \pmod{64}, \lambda \in \{2, 3\},$$

and the reverse implication follows from the previous analysis. If $\alpha \geq 3$, the previous analysis also concludes. \square

A.3. Definition of $E_{2,\alpha,\beta}$. We are now in position to describe the sets $E_{2,\alpha,\beta}$. We first define the clopen sets

$$A_0 = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \gamma \neq I_2 \pmod 2\},$$

$$A_1 = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \gamma = I_2 \pmod 2, \gamma \neq \pm I_2 \pmod 4\},$$

$$A_2 = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \gamma = \pm I_2 \pmod 4, \gamma \neq \pm I_2, \pm 3I_2 \pmod 8\},$$

and for $\alpha \geq 3$,

$$A_\alpha = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \gamma = \pm I_2 \pmod{2^\alpha}, \gamma \neq \pm I_2, \pm(1 + 2^\alpha I_2) \pmod{2^{\alpha+1}}\},$$

$$B_\alpha = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \gamma = \pm(1 + 2^{\alpha-1})I_2 \pmod{2^\alpha}\},$$

with the convention that $B_0 = B_1 = B_2 = \emptyset$. Thus by Lemmata 6.1 and A.1, for any $\gamma \in \Gamma$ hyperbolic,

$$v_2(u_\gamma) = \alpha \iff \gamma \in A_\alpha \sqcup B_\alpha.$$

The sets A_k, B_k form a partition of $\mathrm{SL}(2, \mathbb{Z}_p) - \{\pm I_2\}$.

Similarly, we put:

$$C_0 = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \mathrm{Tr}(\gamma) \pmod{16} \notin \{2, 6, 10, 14\}\},$$

$$C_1 = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \mathrm{Tr}(\gamma) \pmod{16} \in \{6, 10\}\},$$

$$C_2 = \{\gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \mathrm{Tr}(\gamma) \pmod{64} \in \{14, 30, 34, 50\}\},$$

and for $k \geq 3$,

$$C_k = \left\{ \gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \mathrm{Tr}(\gamma) = \pm(2 + \lambda 2^{2k}) \bmod 2^{2k+2}, \lambda \in \{2, 3\} \right\}.$$

$$D_k = \left\{ \gamma \in \mathrm{SL}(2, \mathbb{Z}_2) : \mathrm{Tr}(\gamma) = \pm(2 + 2^{2k-2}) \bmod 2^{2k} \right\}$$

By Lemma A.3, for γ hyperbolic,

$$v_2(u_\gamma) + v_2(f_\gamma) = k \iff \gamma \in C_k \sqcup D_k.$$

The sets C_k, D_k form a partition of $\mathrm{SL}(2, \mathbb{Z}_p) - \mathrm{Tr}^{-1}(\pm 2)$. Finally, we put

$$E_{2,\alpha,\beta} = (A_\alpha \sqcup B_\alpha) \cap (C_{\alpha+\beta} \sqcup D_{\alpha+\beta}).$$

So for γ hyperbolic,

$$v_2(u_\gamma) = \alpha, v_2(f_\gamma) = \beta \iff \gamma \in E_{2,\alpha,\beta}.$$

By Lemma A.2, we have for $\alpha \geq 3$, $B_\alpha \subset D_\alpha$. This imply that $B_\alpha \cap C_{\alpha+\beta} = \emptyset$ for all β . So $E_{2,\alpha,\beta}$ admits a slightly simpler expression:

$$E_{2,\alpha,\beta} = \begin{cases} (A_\alpha \cap C_\alpha) \sqcup B_\alpha & \text{if } \beta = 0, \\ A_\alpha \cap (C_{\alpha+\beta} \sqcup D_{\alpha+\beta}) & \text{if } \beta > 0. \end{cases}$$

A.4. The measure of $E_{2,\alpha,\beta}$. The goal of this section is to explain how to get the following array.

Proposition A.4. *The value of $\mu_2(E_{2,\alpha,\beta})$ is given by the following array.*

| $\mu_2(E_{2,\alpha,\beta})$ | $\alpha = 0$ | $\alpha = 1$ | $\alpha = 2$ | $\alpha \geq 3$ |
|-----------------------------|------------------------|------------------------|------------------------|----------------------------|
| $\beta = 0$ | 7/12 | 1/16 | 1/64 | $11/(3 \cdot 2^{3\alpha})$ |
| $\beta = 1$ | 1/8 | 1/32 | $3/2^{3\alpha+2\beta}$ | $3/2^{3\alpha+2\beta}$ |
| $\beta = 2$ | 1/16 | $3/2^{3\alpha+2\beta}$ | $3/2^{3\alpha+2\beta}$ | $3/2^{3\alpha+2\beta}$ |
| $\beta \geq 3$ | $3/2^{3\alpha+2\beta}$ | $3/2^{3\alpha+2\beta}$ | $3/2^{3\alpha+2\beta}$ | $3/2^{3\alpha+2\beta}$ |

For the six cases $\alpha + \beta \leq 2$, these measures can be determined by a computer enumeration of the 196608 matrices in $\mathrm{SL}(2, \mathbb{Z}/64\mathbb{Z})$, as by Lemma A.3, these clopen sets $E_{2,\alpha,\beta}$ are defined modulo 64. This being done, it remains to show the formulas in the two cases $\alpha \geq 3, \beta = 0$ and $\alpha + \beta \geq 3, \beta > 0$. We first compute the number of matrices which correspond to subsets of $(A_\alpha \sqcup B_{\alpha+1}) \cap C_k$ or $(A_\alpha \sqcup B_{\alpha+1}) \cap D_{k+1}$:

Lemma A.5. *Let $k \geq \alpha \geq 0$, with $k \geq 2$. Let $\lambda \in \{1, 2, 3\}$. The number of matrices in $\mathrm{SL}(2, \mathbb{Z}/2^{2k+2}\mathbb{Z})$ that satisfy the 3 conditions*

$$\begin{cases} \gamma = \pm I_2 \bmod 2^\alpha, \\ \gamma \neq \pm I_2 \bmod 2^{\alpha+1}, \\ \mathrm{Tr}(\gamma) = \pm(2 + \lambda 2^{2k}) \bmod 2^{2k+2}, \end{cases}$$

is

$$\begin{cases} 5 \cdot 2^{3k+3} & \text{if } k = \alpha, \lambda = 1 \\ 3 \cdot 2^{3k+3} & \text{if } k = \alpha, \lambda = 2 \text{ or } 3, \\ 3 \cdot 2^{4k-\alpha+3} & \text{otherwise.} \end{cases}$$

Proof. For a matrix γ satisfying the above condition, its trace is $\pm 2 \bmod 2^{2k}$. Since $k \geq 2$, $-2 \not\equiv +2 \bmod 2^{2k}$ so the sign $\epsilon \in \{\pm 1\}$ such that

$$\text{Tr}(\gamma) = \epsilon 2 \bmod 2^{2k},$$

is defined without ambiguity. If $\alpha \geq 2$, the sign \pm in the first condition:

$$\gamma = \pm I_2 \bmod 2^\alpha,$$

is well-defined, and considering the trace $\bmod 2^\alpha$, must agree with the above ϵ . Otherwise, this sign is irrelevant and can be chosen to agree with the above ϵ . So in all cases we are left to count, given a sign $\epsilon \in \{\pm 1\}$, the number of matrices satisfying

$$\begin{cases} \gamma = \epsilon I_2 \bmod 2^\alpha, \\ \gamma \neq \epsilon I_2 \bmod 2^{\alpha+1}, \\ \text{Tr}(\gamma) = \epsilon(2 + \lambda 2^{2k}) \bmod 2^{2k+2}, \\ \det(\gamma) = 1 \end{cases}$$

The first condition may be written as the existence of a, b, c, d in $\mathbb{Z}/2^{2k+2-\alpha}\mathbb{Z}$ such that

$$(A.4) \quad \gamma = \epsilon \left(I_2 + 2^\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right),$$

and the second condition states that not all of them are divisible by 2. By the third condition on the trace,

$$d = \lambda 2^{2k-\alpha} - a.$$

Since $k \geq \alpha$ and $k \geq 2$, by the above equation, d and a have the same parity; in particular the condition that not all a, b, c, d are even is satisfied iff not all of a, b, c are even.

We wish to rewrite the condition that γ has determinant 1, as an equation on a, b, c , assuming that d is given as above.

$$(1 + a2^\alpha)(1 + \lambda 2^{2k} - a2^\alpha) - bc2^{2\alpha} = 1 \bmod 2^{2k+2},$$

which is equivalent to

$$(A.5) \quad (1 + a2^\alpha)\lambda 2^{2(k-\alpha)} - a^2 - bc = 0 \bmod 2^{2(k-\alpha)+2}.$$

First case: $k = \alpha$, so here $\alpha \geq 2$. Then Equation (A.5) is equivalent to:

$$\lambda - a^2 - bc = 0 \pmod{4}.$$

We can count (by computer, or by hand) the number of solution $(x, y, z) \in (\mathbb{Z}/4\mathbb{Z})^3$ to this equation $x^2 + yz = \lambda \pmod{4}$ depending on λ , adding the condition that x, y, z cannot all be even: there are 20 solutions if $\lambda = 1$, and 12 if $\lambda = 2$ or 3. However, a, b, c are in fact elements of $\mathbb{Z}/2^{k+2}\mathbb{Z}$ whose reduction mod 4 are a solution (x, y, z) , so there are $20 \cdot 2^{3k}$ or $12 \cdot 2^{3k}$ solutions (a, b, c) depending on $\lambda = 1$ or 2, 3, for a given sign ϵ . All in all, this gives us $5 \cdot 2^{3k+3}$ or $3 \cdot 2^{3k+3}$ solutions.

Second case: $k > \alpha$. We put $\beta = k - \alpha > 0$. Then, by (A.5), (a, b, c) has to be congruent mod $2^{2\beta}$ to a solution, say a_0, b_0, c_0 , of

$$(A.6) \quad X^2 + YZ = 0 \pmod{2^{2\beta}},$$

with the additional requirement that a_0, b_0, c_0 cannot be all even. With the help of Lemma 7.2, one can show that there are $2^3(2^{4\beta-4} + 2^{4\beta-5} - 2^{3\beta-4})$ solutions of (A.6) for which a_0, b_0, c_0 are all even, among the $2^{4\beta} + 2^{4\beta-1} - 2^{3\beta-1}$ general solutions of (A.6). This means that $(a_0, b_0, c_0) \in \mathbb{Z}/2^{2\beta}\mathbb{Z}$ are freely chosen in a set of cardinality

$$3 \cdot 2^{4\beta-2}.$$

We may now define again auxiliary variables $(u, v, w) \in (\mathbb{Z}/4\mathbb{Z})^3$ such that

$$a = a'_0 + u2^{2\beta}, \quad b = b'_0 + v2^{2\beta}, \quad c = c'_0 + w2^{2\beta},$$

with a'_0, b'_0, c'_0 being fixed representatives of a_0, b_0, c_0 in $\mathbb{Z}/2^{2\beta+2}\mathbb{Z}$. By construction, there exists some $\nu \in \mathbb{Z}/4\mathbb{Z}$ such that $a_0'^2 + b_0'c_0' = \nu 2^{2\beta}$.

Equation (A.5) can be rewritten as an equation on u, v, w :

$$\lambda(1 + a'_0 2^\alpha) 2^{2\beta} = \nu 2^{2\beta} + 2^{2\beta}(2a'_0 u + b'_0 w + c'_0 v) \pmod{2^{2\beta+2}},$$

so is equivalent to

$$(A.7) \quad \lambda(1 + a'_0 2^\alpha) - \nu = 2a'_0 u + b'_0 w + c'_0 v \pmod{4},$$

which is linear in (u, v, w) . However, inspecting again the equation $X^2 + YZ = 0$, we see that if a_0 is odd then b_0 and c_0 must also be odd. Since a_0, b_0, c_0 are not all even, at least one of the two numbers b_0 or c_0 must be odd, so the linear map $(\mathbb{Z}/4\mathbb{Z})^3 \rightarrow \mathbb{Z}/4\mathbb{Z}$, $(u, v, w) \mapsto 2a'_0 u + b'_0 w + c'_0 v$ must be a surjective group morphism, in this case a 16-to-1 map. This means that the number of solutions (u, v, w) of (A.7) is 16. Thus, the number of solutions $(a, b, c) \in \mathbb{Z}/2^{2\beta+2}\mathbb{Z}$ of (A.5) is $16 \cdot 3 \cdot 2^{4\beta-2} = 3 \cdot 2^{4\beta+2}$. However, (a, b, c) are in fact chosen in $\mathbb{Z}/2^{k+\beta+2}\mathbb{Z}$, so we must multiply this number by $2^{3\alpha}$. Remembering that we also had to choose a sign ϵ , we get that the number we seek is

$$3 \cdot 2^{4k-\alpha+3}.$$

□

The next two lemma establish the remaining cases of Proposition A.4.

Lemma A.6. *For $\alpha \geq 3$,*

$$\mu_p(E_{2,\alpha,0}) = \frac{11}{3}2^{-3\alpha}$$

Proof. Recall that for $\alpha \geq 3$,

$$E_{2,\alpha,0} = (A_\alpha \cap C_\alpha) \sqcup B_\alpha.$$

The measure of B_α is straightforward to compute: among the $3 \cdot 2^{3\alpha-2}$ matrices of $\text{SL}(2, \mathbb{Z}/2^\alpha\mathbb{Z})$, we must count the two matrices $\pm(1 + 2^{\alpha-1})I_2$. So

$$\mu_2(B_\alpha) = \frac{1}{3 \cdot 2^{3\alpha-3}}.$$

We are left to determine the measure of $A_\alpha \cap C_\alpha$; this is almost the previous lemma for $k = \alpha$ and $\lambda = 2, 3$, except we did not a priori excludes matrices that may be congruent to $\pm(1 + 2^\alpha)I_2 \pmod{2^{\alpha+1}}$. However, such matrices are in $B_{\alpha+1} \subset D_{\alpha+1}$ so never belong to C_k for any k . Thus, summing the number of solution for $\lambda = 2$ and $\lambda = 3$, we get:

$$\mu_2(A_\alpha \cap C_\alpha) = \frac{3 \cdot 2^{3\alpha+3} + 3 \cdot 2^{3\alpha+3}}{3 \cdot 2^{6\alpha+4}} = \frac{1}{2^{3\alpha}}.$$

So we get

$$\mu_2(E_{2,\alpha,0}) = \frac{11}{3 \cdot 2^{3\alpha}}. \quad \square$$

Lemma A.7. *For $\beta \geq 1$ and $\alpha + \beta \geq 3$,*

$$\mu_2(E_{2,\alpha,\beta}) = \frac{3}{2^{3\alpha+2\beta}}.$$

Proof. Lemma A.5 with $\lambda = 2, 3$ and $k = \alpha + \beta$ tells us how to compute the measure of $(A_\alpha \sqcup B_{\alpha+1}) \cap C_{\alpha+\beta} = A_\alpha \cap C_{\alpha+\beta}$. Adding the cases $\lambda = 2, 3$, we get that for $\beta > 0$,

$$\mu_2(A_\alpha \cap C_{\alpha+\beta}) = \frac{6 \cdot 2^{3\alpha+4\beta+3}}{3 \cdot 2^{3(2\alpha+2\beta+2)-2}} = \frac{1}{2^{3\alpha+2\beta}}.$$

We now make a distinction between the cases $\beta > 1$ and $\beta = 1$. First assume that $\beta > 1$. Lemma A.5 with $\lambda = 1$, $k = \alpha + \beta - 1$ gives us

$$\mu_2((A_\alpha \sqcup B_{\alpha+1}) \cap D_{\alpha+\beta}) = \frac{3 \cdot 2^{3\alpha+4\beta-1}}{3 \cdot 2^{3(2\alpha+2\beta)-2}} = \frac{2}{2^{3\alpha+2\beta}},$$

where we used the last formula of Lemma A.5 since $k > \alpha$. Also, since $\beta > 1$, $B_{\alpha+1} \cap D_{\alpha+\beta} = \emptyset$ since $B_{\alpha+1} \subset D_{\alpha+1}$, so

$$\mu_2(E_{2,\alpha,\beta}) = \mu_2(A_\alpha \cap C_{\alpha+\beta}) + \mu_2(A_\alpha \cap D_{\alpha+\beta}) = \frac{3}{2^{3\alpha+2\beta}},$$

as claimed.

We now treat the case where $\beta = 1$. Then Lemma A.5 with $\lambda = 1$, $k = \alpha$ gives us

$$\mu_2((A_\alpha \sqcup B_{\alpha+1}) \cap D_{\alpha+1}) = \frac{5 \cdot 2^{3\alpha+3}}{3 \cdot 2^{3(2\alpha+2)-2}} = \frac{5}{3 \cdot 2^{3\alpha+1}},$$

Since $\beta = 1$, we have

$$(A_\alpha \sqcup B_{\alpha+1}) \cap D_{\alpha+1} = (A_\alpha \cap D_{\alpha+1}) \sqcup B_{\alpha+1},$$

so

$$\mu_2(A_\alpha \cap D_{\alpha+1}) = \frac{5}{3 \cdot 2^{3\alpha+1}} - \frac{1}{3 \cdot 2^{3\alpha}} = \frac{1}{2^{3\alpha+1}},$$

so using $E_{2,\alpha,1} = (A_\alpha \cap D_{\alpha+1}) \sqcup (A_\alpha \cap C_{\alpha+1})$, the result is

$$\mu_2(E_{2,\alpha,1}) = \frac{1}{2^{3\alpha+1}} + \frac{1}{2^{3\alpha+2}} = \frac{3}{2^{3\alpha+2\beta}}. \quad \square$$

Appendix B. Enumeration of geodesics by length

Here we wish to discuss in more depth the relationship between picking randomly a periodic geodesic on the modular surface using the uniform measure on those of length $\leq L$, and picking a matrix at random using \mathbb{P}_T , which is likely hyperbolic and primitive, and then considering its conjugacy class in $\mathrm{PSL}(2, \mathbb{Z})$, which then correspond to a random periodic geodesic. The point is that these two probabilities are, in some sense, absolutely continuous to each other.

Let E_Γ be a subset of Γ that is a union of conjugacy classes of primitive hyperbolic elements, symmetric in the sense that if $\gamma \in E_\Gamma$, then $-\gamma \in E_\Gamma$. Then E_Γ correspond to a subset $E_\mathcal{G}$ of the set of oriented, primitive periodic geodesic on the modular surface. Denote by \mathcal{G}_L the subset of oriented, primitive periodic geodesic on the modular surface of length $\leq L$.

Theorem B.1. *Let $E_\Gamma \subset \Gamma$, $E_\mathcal{G}$ be as above. Then E_Γ has zero natural density if and only if $E_\mathcal{G}$ has zero natural density. More precisely,*

$$\left(\limsup_{T \rightarrow +\infty} \frac{|E_\Gamma \cap \Gamma_T|}{|\Gamma_T|} > 0 \right) \iff \left(\limsup_{L \rightarrow +\infty} \frac{|E_\mathcal{G} \cap \mathcal{G}_L|}{|\mathcal{G}_L|} > 0 \right).$$

Thus properties of conjugacy classes which are true for almost all (resp. almost no) matrices of $\mathrm{SL}(2, \mathbb{Z})$, ordered by norm, are also true for almost all (resp. almost no) of the corresponding periodic geodesics, ordered by length. The rest of the section is devoted to proving Theorem B.1.

B.1. Enumeration by norm and displacement. It is first worth a reminder that the enumeration by norm is nothing else than a enumeration by displacement. Indeed, consider the point $i \in \mathbb{H}^2$ in the upper half-plane

model of the hyperbolic plane, then the Frobenius norm of γ and the hyperbolic distance between i and γi are related by (see for example [19, 3.1])

$$\|\gamma\|^2 = 2 \cosh d_{\mathbb{H}^2}(\gamma i, i),$$

where $\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\| = \sqrt{a^2 + b^2 + c^2 + d^2}$ is the Frobenius norm. So if T and R are related by $T^2 = 2 \cosh R$, then

$$\Gamma_T = \{\gamma \in \Gamma : d_{\mathbb{H}^2}(\gamma i, i) \leq R\}.$$

The following lemma is classical:

Lemma B.2. *Let $\gamma \in \Gamma$ be hyperbolic, denote by $\mathcal{A}_\gamma \subset \mathbb{H}^2$ be the axis of the hyperbolic transformation γ acting on \mathbb{H}^2 , and ℓ_γ its translation length. Let $c = -2 \log \frac{3\sqrt{5}-5}{4} > 0$, then*

$$d_{\mathbb{H}^2}(\gamma i, i) \geq 2d_{\mathbb{H}^2}(i, \mathcal{A}_\gamma) + \ell_\gamma - c.$$

Proof. Let z be the orthogonal projection of i onto the axis \mathcal{A}_γ , and w the midpoint of the geodesic segment $[z, \gamma z]$. Note that the orthogonal projection of γi onto \mathcal{A}_γ is γz and that the orthogonal to \mathcal{A}_γ through w cuts the segment $[i, \gamma i]$ in its midpoint, which we call v . Thus the quadrilateral given by i, v, w, z is right-angled at the three vertices v, w, z , one may recognize a *Lambert quadrilateral*. By [9, Theorem 32.21],

$$\sinh d_{\mathbb{H}^2}(i, v) = \sinh d_{\mathbb{H}^2}(w, z) \cdot \cosh d_{\mathbb{H}^2}(i, z),$$

that is

$$\sinh \frac{d_{\mathbb{H}^2}(i, \gamma i)}{2} = \sinh \frac{\ell_\gamma}{2} \cdot \cosh d_{\mathbb{H}^2}(i, \mathcal{A}_\gamma).$$

Since $\gamma \in \Gamma$ is hyperbolic, $\ell_\gamma \geq 2 \log \frac{3+\sqrt{5}}{2}$ (the systole of the surface $\Gamma \backslash \mathbb{H}^2$), so we have $\sinh \frac{\ell_\gamma}{2} = \frac{1}{2} e^{\ell_\gamma/2} (1 - e^{-\ell_\gamma}) \geq \frac{3\sqrt{5}-5}{4} e^{\ell_\gamma/2}$. Thus

$$\frac{1}{2} e^{d_{\mathbb{H}^2}(i, \gamma i)/2} \geq \frac{3\sqrt{5}-5}{4} e^{\ell_\gamma/2} \cdot \frac{1}{2} e^{d_{\mathbb{H}^2}(i, z)}$$

Taking the logarithm concludes. \square

We now prove the implication \Rightarrow . Let E_Γ be a set of positive upper density, say ϵ , for the enumeration by norm. The main point is to bound how many times we can pick different elements of Γ_T corresponding to the same periodic geodesic, linearly in the length.

By the formula (1.3), we can find a $\beta = \beta_\epsilon > 0$ such that

$$\limsup_{T \rightarrow +\infty} \mathbb{P}_T(\gamma \in E_\Gamma, 2 \log(T) \geq \ell_\gamma \geq 2 \log(T) - \beta) > \epsilon/2.$$

Denote by $E'_T = \{\gamma \in E_\Gamma : 2 \log T \geq \ell_\gamma \geq 2 \log(T) - \beta\}$. For any $\gamma \in E'_T$, by the lemma,

$$d_{\mathbb{H}^2}(i, \mathcal{A}_\gamma) \leq \frac{1}{2}(R(T) - 2 \log(T) + \beta + c),$$

where $R(T) = \cosh^{-1}(T^2/2) = 2\log(T) + o(1)$. Thus there is a uniform bound $r = r_\epsilon > 0$ on the distance between i and the axis \mathcal{A}_γ of γ for all elements of E'_T . Pick now some $\gamma_0 \in E'_T$, and let m_{γ_0} be the number of $\gamma \in E'_T$ which are conjugate to γ_0 or $-\gamma_0$. Consider the following picture: draw a disk of radius $r + 1$ centered on i , and its intersection with the axis \mathcal{A}_γ for all $\gamma \in E'_T$ conjugate to $\pm\gamma_0$. What we see is a family of m_{γ_0} geodesic segments. Since the closest point to i of \mathcal{A}_γ is at distance $\leq r$, these segments must intersect the disk of radius r , and their endpoints belong to the circle of radius $r + 1$, so all segments have length at least 2. Let $N = N_\epsilon$ be the number of fundamental domains for Γ intersecting the disk of radius $r + 1$. Then if we project the picture to $\Gamma \backslash \mathbb{H}^2$, comparing the total length of the segments involved, we see that

$$2m_{\gamma_0} \leq N\ell_{\gamma_0}.$$

This is the required linear bound on m_{γ_0} in terms of ℓ_{γ_0} . Thus, if for some large $T > 0$, $\mathbb{P}_T(E'_T) > \epsilon/4$, meaning there are at least ϵT^2 matrices in E'_T , that are representing at least $\frac{2}{N}\epsilon T^2/(2\log(T))$ distinct conjugacy classes. Now put $L = 2\log(T)$, using the well-known asymptotic

$$|\mathcal{G}_L| \sim \frac{e^L}{L} = \frac{T^2}{2\log(T)},$$

we conclude that $E_{\mathcal{G}}$ must have upper density at least $\frac{2}{N}\epsilon$.

B.2. Enumeration by length. We study the converse implication. Assume that $E_{\mathcal{G}}$ has positive density. This time, the point is that for most geodesics, one can construct a number of lifts close to i which is at least linear in the length. For an oriented periodic geodesic $\lambda \in \mathcal{G}_L$, denote by μ_λ the length measure on λ , normalised to be a probability.

Lemma B.3. *There exists a compact set $C \subset G$ and a subset $E'_G \subset E_G$ of the same upper density, such that for every $\lambda \in E'_G$, $\mu_\lambda(C) > 1/3$.*

Proof. Recall that periodic geodesics on the modular surface equidistribute (see for example [16, Theorem 1], or [20]), meaning the following average converges:

$$\lim_{L \rightarrow +\infty} \frac{1}{|\mathcal{G}_L|} \sum_{\lambda \in \mathcal{G}_L} \mu_\lambda = \mu_{\Gamma \backslash G}.$$

Since Bowen, equidistribution of periodic orbits is often expressed as a convergence in average, but in fact, together with ergodicity, it implies that after removing a set of periodic geodesics of zero density, μ_λ must be close to $\mu_{\Gamma \backslash G}$. Let's sketch a proof: given a continuous function $f : \Gamma \backslash G \rightarrow \mathbb{R}$ with compact support and $\eta > 0$, define

$$\mathcal{G}_{bad} = \left\{ \lambda : \mu_\lambda(f) \leq \mu_{\Gamma \backslash G}(f) - \eta \right\}.$$

We wish to show that the above bad set of geodesics must be of zero density. This will be sufficient by a diagonal argument using a countable dense subset of such functions f and rational $\eta > 0$. We proceed by contradiction.

Let $\mathcal{G}_{bad,L} = \mathcal{G}_{bad} \cap \mathcal{G}_L$, and assume there exists a sequence $L_n \rightarrow +\infty$ such that

$$\alpha := \lim_n \frac{|\mathcal{G}_{bad,L_n}|}{|\mathcal{G}_{L_n}|} > 0.$$

Write $\mu_{bad,L} = \frac{1}{|\mathcal{G}_{bad,L}|} \sum_{\lambda \in \mathcal{G}_{bad,L}} \mu_\lambda$, and $\mu'_L = \frac{1}{|\mathcal{G}_L - \mathcal{G}_{bad,L}|} \sum_{\lambda \in \mathcal{G}_L - \mathcal{G}_{bad,L}} \mu_\lambda$.

By compactness of the set of Radon measures on $\Gamma \backslash G$ of total mass ≤ 1 , we can extract a subsequence, still denoted by $(L_n)_n$, such that μ_{bad,L_n} converges to some measure ν and μ'_{L_n} to some measure ν' of mass ≤ 1 . Notice that both $\mu_{bad,L}$ and μ'_L are invariant by the geodesic flow, so ν and ν' are also invariant measures, and $\nu(f) \leq \mu_{\Gamma \backslash G}(f) - \eta$ by construction, so $\nu \neq \mu_{\Gamma \backslash G}$.

By equidistribution in average, we have

$$\lim_{n \rightarrow \infty} \frac{|\mathcal{G}_{bad,L_n}|}{|\mathcal{G}_{L_n}|} \mu_{bad,L_n} + \left(1 - \frac{|\mathcal{G}_{bad,L_n}|}{|\mathcal{G}_{L_n}|}\right) \mu'_{L_n} = \mu_{\Gamma \backslash G},$$

so

$$\alpha \nu + (1 - \alpha) \nu' = \mu_{\Gamma \backslash G},$$

so both ν, ν' must be probabilities, and the above equality contradicts the ergodicity of $\mu_{\Gamma \backslash G}$ by the geodesic flow.

Returning to our problem, choose a compact subset $C \subset \Gamma \backslash G$ of Haar measure bigger than $\mu_{\Gamma \backslash G}(C) \geq 1/2$, of negligible boundary $\mu_{\Gamma \backslash G}(\partial C) = 0$. Approximate the characteristic function of C from below by a non-negative continuous function f with compact support and Haar integral $> 5/12$. The previous argument with $\eta = 1/12$ gives us a set \mathcal{G}_{bad} of zero density such that for all $\lambda \notin \mathcal{G}_{bad}$, the geodesic λ spends at least $1/3$ of its time in C , so $E'_\mathcal{G} = E_\mathcal{G} - \mathcal{G}_{bad}$ satisfies the required property. \square

Let \tilde{C} be a compact set in G surjecting onto C . Let M be the diameter of \tilde{C} . For such a geodesic $\lambda \in E'_\mathcal{G}$, one can pick vectors v_1, \dots, v_k on λ , with $v_i \in C$, such that v_i, v_j are images of each other by the geodesic flow of time at least $M + 1$ for $i \neq j$. It is possible to do this with $k \geq \ell(\lambda)/3(M + 1)$. Pick now lifts $\tilde{v}_i \in \tilde{C}$ of each v_i ; the geodesics in \mathbb{H}^2 defined by the vectors \tilde{v}_i are all lifts of λ , at bounded distance from the origin i . They are distinct since otherwise, \tilde{v}_i and \tilde{v}_j would be at distance $\leq M$ on the same geodesic, so one could flow from one to the other in time $< M$.

Thus, for each geodesic λ of length ℓ_λ that spend a proportion of time $\geq 1/3$ in C , one could construct $c \cdot \ell$ different lifts at bounded distance from

the point i , for some $c > 0$. Now pick some large L such that

$$\frac{|E'_G \cap \mathcal{G}_L|}{|\mathcal{G}_L|} > \epsilon.$$

By exponential growth, there exists $\eta > 0$ such that the set

$$E'_L = \{\lambda \in E'_G \cap \mathcal{G}_L : \ell_\lambda > \eta L\},$$

satisfies $|E'_L| > \epsilon |\mathcal{G}_L|/2 \geq \epsilon e^L/3L$, and each of these geodesics has at least $c\eta L$ different lifts at bounded distance from i . For each of those $c\eta e^L$ lift, there exists an hyperbolic transformation $\gamma \in \Gamma$ whose axis is precisely the lift. We have to show that they all lie in Γ_T for some T not far from $e^{L/2}$. By the triangle inequality,

$$d_{\mathbb{H}^2}(\gamma i, i) \leq \ell_\gamma + 2d_{\mathbb{H}^2}(i, \mathcal{A}_\gamma) \leq L + O(1),$$

so we have $\gamma \in \Gamma_T$ with $T < Ce^{L/2}$, for some well-chosen $C > 0$.

Acknowledgement

I would like to thank the anonymous referee for useful suggestions and for pointing to me the references [7, 17, 18].

References

- [1] M. BHARGAVA, A. SHANKAR & X. WANG, “Squarefree values of polynomial discriminants I”, *Invent. Math.* **228** (2022), no. 3, p. 1037-1073.
- [2] ———, “Squarefree values of polynomial discriminants II”, 2022, <https://arxiv.org/abs/2207.05592>.
- [3] J. BOURGAIN & A. KONTOROVICH, “Beyond expansion II: low-lying fundamental geodesics”, *J. Eur. Math. Soc.* **19** (2017), no. 5, p. 1331-1359.
- [4] K. CONRAD, “The conductor ideal”, <https://kconrad.math.uconn.edu/blurbs/gradnumthy/conductor.pdf>.
- [5] M. EINSIEDLER, E. LINDENSTRAUSS, P. MICHEL & A. VENKATESH, “The distribution of closed geodesics on the modular surface, and Duke’s theorem”, *Enseign. Math. (2)* **58** (2012), no. 3-4, p. 249-313.
- [6] G. H. HARDY & E. M. WRIGHT, *An introduction to the theory of numbers*, fifth ed., Clarendon Press, 1979, xvi+426 pages.
- [7] Y. HASHIMOTO, “Asymptotic formulas for class number sums of indefinite binary quadratic forms on arithmetic progressions”, *Int. J. Number Theory* **9** (2013), no. 1, p. 27-51.
- [8] W. LUO, Z. RUDNICK & P. SARNAK, “The variance of arithmetic measures associated to closed geodesics on the modular surface”, *J. Mod. Dyn.* **3** (2009), no. 2, p. 271-309.
- [9] G. E. MARTIN, *The foundations of geometry and the non-Euclidean plane.*, corr. 4th printing ed., Undergraduate Texts in Mathematics, Springer, 1998.
- [10] F. MAUCOURANT, “Homogeneous asymptotic limits of Haar measures of semisimple linear groups and their lattices”, *Duke Math. J.* **136** (2007), no. 2, p. 357-399.
- [11] C. MCMULLEN, “Ergodic Theory, Geometry and Dynamics”, 2020, <https://people.math.harvard.edu/~ctm/papers/home/text/class/notes/ergodic/course.pdf>.
- [12] T. MIYAKE, *Modular forms. Transl. from the Japanese by Yoshitaku Maeda*, corrected 2nd printing ed., Springer Monographs in Mathematics, Springer, 2006.
- [13] C. J. MOZZOCHI, “Linking numbers of modular geodesics”, *Isr. J. Math.* **195** (2013), p. 71-95.
- [14] J. NEUKIRCH & N. SCHAPPACHER, *Algebraic Number Theory*, Grundlehren der Mathematischen Wissenschaften, vol. 322, Springer, 2013.

- [15] A. NEVO & P. SARNAK, “Prime and almost prime integral points on principal homogeneous spaces”, *Acta Math.* **205** (2010), no. 2, p. 361-402.
- [16] M. POLLICOTT, “Distribution of closed geodesics on the modular surface and quadratic irrationals”, *Bull. Soc. Math. Fr.* **114** (1986), p. 431-446.
- [17] N. RAULF, “Asymptotics of class numbers for progressions and for fundamental discriminants”, *Forum Math.* **21** (2009), no. 2, p. 221-257.
- [18] ———, “Limit distribution of class numbers for discriminants in progressions and fundamental discriminants”, *Int. J. Number Theory* **12** (2016), no. 5, p. 1237-1258.
- [19] I. RIVIN, “How to pick a random integer matrix? (and other questions)”, *Math. Comput.* **85** (2016), no. 298, p. 783-797.
- [20] T. ROBLIN, “Ergodicité et équidistribution en courbure négative”, *Mém. Soc. Math. Fr., Nouv. Sér.* **95** (2003), p. 1-96.
- [21] C. ROETTGER, “Counting invertible matrices and uniform distribution”, *J. Théor. Nombres Bordeaux* **17** (2005), no. 1, p. 301-322.
- [22] P. SARNAK, “Class numbers of indefinite binary quadratic forms”, *J. Number Theory* **15** (1982), no. 2, p. 229-247.
- [23] ———, “Class numbers of indefinite binary quadratic forms. II”, *J. Number Theory* **21** (1985), p. 333-346.

François MAUCOURANT

Université Rennes, IRMAR UMR 6625, Campus de Beaulieu

35042 Rennes cedex, France

E-mail: francois.maucourant@univ-rennes.fr