QUADRATIC FORMS AND DUKE'S THEOREM

CONSTANTIN KOGLER

Abstract. These notes are a continuation of [Kog19] and we relate quadratic forms to the collection of curves discussed in [Kog19]. We discuss content from Section 3 of the 2012 paper [ELMV12] by Einsiedler, Lindenstrauss, Michel and Venkatesh.

Contents

1. Introduction

We study binary integral quadratic forms

$$Q(X,Y) = aX^2 + bXY + cY^2$$

for $a,b,c \in \mathbb{Z}$. For $\mathbf{X} = (X,Y)^T$, the above quadratic from can we rewritten as

$$Q(X,Y) = \mathbf{X}^T B_Q \mathbf{X} \quad \text{for} \quad B_Q = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}.$$

Moreover, the invariant $d_Q = b^2 - 4ab = -4\det(B_Q)$ is called the discriminant of Q. We note that only the integers $\equiv 0,1 \mod 4$ can be discriminants.

In these notes only primitive quadratic forms are considered, i.e. it is assumed that $\gcd(a,b,c) = 1$. Moreover, we will not distinguish between the quadratic form $Q(X,Y)$ and the tuple $(a,b,c) \in \mathbb{Z}^3$ it defines. We write $\mathcal{Q}_d$ for the set of primitive quadratic forms and only the case where d is positive and not a perfect square is studied.

Recall that $X = \Gamma\backslash G = \mathrm{PGL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2(\mathbb{R})$ can be viewed as the space of lattices up to scaling. We denote by A the diagonal subgroup of $\mathrm{PGL}_2(\mathbb{R})$. As d is a discriminant, we have that $\mathcal{O}_d := \mathbb{Z}[\frac{d+\sqrt{d}}{2}] \subset K := \mathbb{Q}(\sqrt{d})$ is an order in K. Recall that

$$\mathrm{Pic}(\mathcal{O}_d) = \{\text{invertible fractional } \mathcal{O}_d\text{-ideal of } K\}/K^\times.$$

In [Kog19] we related to each invertible fractional $\mathcal{O}_d$-ideal of K a periodic A-orbit $x_{\mathfrak{a}}.A$, which only depends on the equivalence class $[\mathfrak{a}] \in \mathrm{Pic}(\mathcal{O}_d)$. This allows us to define

$$\mathcal{G}_d = \bigcup_{[\mathfrak{a}]\in\mathrm{Pic}(\mathcal{O}_d)} x_{\mathfrak{a}}.A. \tag{1.1}$$

The set $\mathcal{G}_d$ supports a canonical A-invariant measure, which we denote by μ_d.

A first aim of these notes is to relate the collection of periodic A-orbits $\mathcal{G}_d$ to quadratic forms. More precisely, we will describe $\mathcal{G}_d$ in terms of information given by quadratic forms. This will be done in Section 3.

Recall that the height of a lattice L is defined as

$$\mathrm{ht}(L) = \left(\frac{\min_{v\in L\setminus\{0\}} ||v||}{\mathrm{vol}(L)^{1/2}}\right)^{-1}.$$

We write $X_{\geq H} = \{x \in X : \mathrm{ht}(x) \geq H\}$ and $X_{<H} = \{x \in X : \mathrm{ht}(x) < H\}$. Proposition 3.3. of [ELMV12] (and Proposition 7.4 of [Kog19]) implies that for any $\varepsilon > 0$,

$$\mu_d(X_{\geq d^\varepsilon}) \longrightarrow 0$$

and $d \to \infty$.

With the help the viewpoint on $\mathcal{G}_d$ given by quadratic forms, we will be able to estimate how many of the points $(x_1, x_2) \in (\mathcal{G}_d \cap X_{\leq H})^2$ are δ-close. To make this more precise, we next discuss how to define a metric on X. First recall that there is a left-invariant metric d_G on G (see Section 9.3 of [EW11]), i.e. a metric that induces the topology of G and satisfies $d_G(hg_1, hg_2) = d_G(g_1, g_2)$ for any $g_1, g_2, h \in G$. This metric induces a well-defined metric on $X = \Gamma\backslash G$ given as

$$d_X(\Gamma g_1, \Gamma g_2) = \inf_{\gamma\in\Gamma} d_G(g_1, \gamma g_2).$$

We now can state the main proposition of these notes.

Proposition 1.1. *(Main Proposition, Proposition 3.6. of* [ELMV12]*) For any $\varepsilon > 0$ and H large we have*

$$(\mu_d \times \mu_d)\left(\{(x, y) \in X_{<H} \times X_{<H} : d_X(x, y) < \delta\}\right) \ll_\varepsilon H^4\delta^3 d^\varepsilon,$$

where $d^{-1/4} \leq \delta \leq H^{-2}$.

1.1. **Acknowledgment.** I thank Menny Aka and Manfred Einsiedler for helping me understand this material. Moreover, I express my gratitude to Manuel Lüthi for a useful remark concerning the action of $\mathrm{GL}_2(\mathbb{R})$ on the space of quadratic forms. Finally, I thank the attendees of my talk on this topic for their presence.

2. Quadratic Forms and the Picard Group

For any $g \in \mathrm{GL}_2(\mathbb{R})$ and any binary quadratic from Q (not necessarily an element of $\mathcal{Q}_d$) we define

$$(g.Q)(X,Y) = \frac{1}{\det(g)} Q((X,Y)g)$$

and we note that

$$B_{g.Q} = \frac{1}{\det(g)} g B_Q g^T.$$

So we see that $g.Q$ has the same discriminant as Q. Thus $\mathrm{GL}_2(\mathbb{R})$ acts on the space of all quadratic forms of discriminant d and the action factors through $\mathrm{PGL}_2(\mathbb{R})$.

As before we write $\mathcal{Q}_d$ for the set of primitive integral binary quadratic forms of discriminant $d > 0$, for d not a perfect square. If $Q \in \mathcal{Q}_d$, then it is not guaranteed that $g.Q \in \mathcal{Q}_d$ for a general element $g \in \mathrm{GL}_2(\mathbb{R})$. However $\mathrm{PGL}_2(\mathbb{Z})$ defines a well defined action on $\mathcal{Q}_d$.

We say that two quadratic forms $Q_1, Q_2 \in \mathcal{Q}_d$ are equivalent if their $\mathrm{PGL}_2(\mathbb{Z})$-orbits are the same. In particular, this means that Q_1 and Q_2 are equivalent if there is some $g \in \mathrm{PGL}_2(\mathbb{Z})$ such that $Q_1 = g.Q_2$. It is clear that this defines an equivalence relation and we denote by $[\mathcal{Q}_d]$ the set of equivalence classes.

Theorem 2.1. *The set $[\mathcal{Q}_d]$ has the same cardinality as the Picard group $\mathrm{Pic}(\mathcal{O}_d)$.*

For a proof of Theorem 2.1 we refer to [Kog]. In these notes we only construct a map from $\mathrm{Pic}(\mathcal{O}_d) \to [\mathcal{Q}_d]$. Namely, for each invertible fractional $\mathcal{O}_d$-ideal $\mathfrak{a}$ and an integral basis a_1, a_2 of $\mathfrak{a}$ we define

$$Q_{a_1,a_2}(X,Y) = \frac{\mathrm{N}(a_1 X + a_2 Y)}{\mathrm{N}(\mathfrak{a})}.$$

Furthermore, $Q_{(a_1,a_2)}$ has discriminant d as

$$\begin{aligned} d_{Q_{(a_1,a_2)}} &= \frac{1}{\mathrm{N}(\mathfrak{a})^2}((a_1\sigma a_2 + \sigma a_1 a_2)^2 - 4a_1\sigma a_1 \sigma a_1 a_2) \\ &= \frac{1}{\mathrm{N}(\mathfrak{a})^2}(a_1\sigma a_2 - \sigma a_1 a_2)^2 \\ &= \frac{1}{\mathrm{N}(\mathfrak{a})^2} \det\left(\begin{pmatrix} a_1 & a_2 \\ \sigma a_1 & \sigma a_2 \end{pmatrix}\right)^2 = \frac{d(\mathfrak{a})}{\mathrm{N}(\mathfrak{a})^2} = d \end{aligned}$$

Finally, $Q_{(a_1,a_2)}$ as an element of $[\mathcal{Q}_d]$ does not depend on the choice of integral basis and so defines an element $Q_\mathfrak{a} \in [\mathcal{Q}_d]$. Moreover, $Q_\mathfrak{a}$ only depends on the ideal class $[\mathfrak{a}] \in \mathrm{Pic}(\mathcal{O}_d)$. So we have constructed a map

$$\mathrm{Pic}(\mathcal{O}_d) \longrightarrow [\mathcal{Q}_d], \qquad [\mathfrak{a}] \mapsto [Q_\mathfrak{a}],$$

which is in fact a bijection.

3. A-orbits associates to Quadratic Forms

Write $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}$. For $(a,b,c) \in \mathcal{Q}_d$ we set

$$x_{a,b,c,\pm} = \frac{-b + \sqrt{d}}{2c}. \in \mathbb{R} = \partial\mathbb{H}.$$

These two points uniquely determine a geodesic on $\mathbb{H}$ and also a lift onto the unit tangent bundle $T^1(\mathbb{H}) \cong \mathrm{SL}(\mathbb{R})$ which we call $\gamma_{(a,b,c)}$. Denote by Q_0 the quadratic form $Q_0(X,Y) = XY$. We claim that

$$\gamma_{(a,b,c)} = \{g \in \mathrm{PGL}_2(\mathbb{R}) : \sqrt{d}(g.Q_0)(X,Y) = aX^2 + bXY + cY^2\}. \quad (3.1)$$

To see this we denote

$$h_{a,b,c} = \begin{pmatrix} b+\sqrt{d} & b-\sqrt{d} \\ 2c & 2c \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Then $wh_{a,b,c} = \left(\begin{smallmatrix} -2c & -2c \\ b+\sqrt{d} & b-\sqrt{d} \end{smallmatrix}\right)$ and so

$$wh_{a,b,c}.0 = -\frac{2c}{b-\sqrt{d}} = -\frac{2c}{b-\sqrt{d}}\frac{b+\sqrt{d}}{b+\sqrt{d}} = \frac{-2cb - 2c\sqrt{d}}{-4ac} = \frac{b+\sqrt{d}}{2a}$$

and

$$wh_{a,b,c}.\infty = -\frac{2c}{b+\sqrt{d}} = -\frac{2c}{b+\sqrt{d}}\frac{b-\sqrt{d}}{b-\sqrt{d}} = \frac{-2cb + 2c\sqrt{d}}{-4ac} = \frac{b-\sqrt{d}}{2a}.$$

Furthermore we calculate

$$\frac{1}{\det(h_{a,b,c})} h_{a,b,c} \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} h^T_{a,b,c} = \frac{1}{4c\sqrt{d}} \begin{pmatrix} b^2-d & 2cb \\ 2cb & 4c^2 \end{pmatrix} = \frac{1}{\sqrt{d}} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$$

and this shows that

$$\sqrt{d}(h_{a,b,c}.Q_0)(X,Y) = aX^2 + bXY + cY^2.$$

Hence (3.1) follows as A-is the stabilizer of Q_0.

We can project $\gamma_{(a,b,c)}$ to the curve $\gamma^X_{(a,b,c)}$ on $X = T^1(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}) \cong \mathrm{PGL}_2(\mathbb{Z})\backslash\mathrm{PGL}_2(\mathbb{R})$. If $(a,b,c) \in \mathcal{Q}_d$ and $(a',b',c') \in \mathcal{Q}_d$ define the same $\mathrm{PGL}_2(\mathbb{Z})$ equivalence class, then $x_{a,b,c,\pm}$ and $x_{a',b',c',\pm}$ lie on the same $\mathrm{PGL}_2(\mathbb{Z})$ orbit. Thus $\gamma^X_{a,b,c}$ only depends on the class $[(a,b,c)] \in [\mathcal{Q}_d]$. Thus we can define

$$\mathcal{G}_d = \bigcup_{[(a,b,c)] \in [\mathcal{Q}_d]} \gamma^X_{[(a,b,c)]}.$$

This definition agrees with (1.1), see [ELMV12] section 2.3 for a discussion why this is the case and [Kog] section 4.4. for a detailed proof.

4. Representations of Quadratic Forms

Let q be an integral quadratic form in n-variables and let Q be one in m-variables, where we assume $n \leq m$. We call a $\mathbb{Z}-$linear map $\iota : \mathbb{Z}^n \to \mathbb{Z}^m$ a **representation** of q by Q if for all $x \in \mathbb{Z}^n$ we have

$$Q(\iota(x)) = q(x).$$

Denote by $R_Q(q)$ the set of such representations. We write

$$\mathrm{SO}_Q(\mathbb{Z}) = \{A \in M_m(\mathbb{Z}) \,:\, Q(Ax) = Q(x) \text{ for all } x \in \mathbb{Z}^m\}$$

for the special orthogonal group with respect to Q. Then SO_Q naturally acts on the set $R_Q(q)$ and the quotient $\mathrm{SO}_Q(\mathbb{Z})\backslash R_Q(q)$ is finite.

For a more extended discussion concerning the representation of integral quadratic forms we refer to Section 3.2 of [ELMV12]. We will be only interested in the representation of the quadratic form $q(X,Y) = dX^2 + \ell XY + dY^2$ for d as above and ℓ some integer by the ternary form given by the discriminant $\mathrm{disc}(X,Y,Z) = Y^2 - 4XZ$. In this situation the following result holds.

Corollary 4.1. *(Corollary 3.5 of* [ELMV12]*) Assume that $\ell \neq \pm 2d$, then*

$$|\mathrm{SO}_{\mathrm{disc}}(\mathbb{Z})\backslash\{(\mathbb{Z}^2, dX^2 + \ell XY + dY^2) \hookrightarrow (\mathbb{Z}^3, \mathrm{disc})\}| \ll_\varepsilon f \max(|d|, |\ell|)^\varepsilon$$

where f^2 is the largest square divisor of $\gcd(d,\ell)$.

Moreover, we next explain how we can embed $\mathrm{PGL}_2(\mathbb{Z})$ into $\mathrm{SO}_{\mathrm{disc}}(\mathbb{Z})$ with finite index. Denote

$$v = v(a,b,c) = \begin{pmatrix} b & 2a \\ -2c & -b \end{pmatrix},$$

so that we can write $\mathfrak{sl}_2(\mathbb{R}) = \{v(a,b,c) \,:\, a,b,c \in \mathbb{R}\}$ and note that $\mathfrak{sl}_2(\mathbb{R})$ is a three dimensional vector space over $\mathbb{R}$ on which we define a quadratic form

$$Q(v(a,b,c)) = \det(v) = -b^2 + 4ac = -\mathrm{disc}(a,b,c).$$

We now view $\mathbb{R}^3 \cong \mathfrak{sl}_2(\mathbb{R})$ and so $(\mathbb{R}^3, Q)$ is a quadratic space. Observe that $\mathrm{PGL}_2(\mathbb{R})$ acts on $\mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ via $g.v = gvg^{-1}$ for $g \in \mathrm{PGL}_2(\mathbb{R})$ and $v \in \mathfrak{sl}_2(\mathbb{R}) \cong \mathbb{R}^3$ and moreover $Q(g.v) = \det(g.v) = \det(v) = Q(v)$. This shows that this action defines an injective representation

$$\mathrm{PGL}_2(\mathbb{R}) \hookrightarrow \mathrm{SO}_Q(\mathbb{R}), \qquad g \mapsto (v \mapsto gvg^{-1})$$

where injectivity follows easily. Thus $\mathrm{PGL}_2(\mathbb{R})$ is isomorphic to either $\mathrm{SO}_Q(\mathbb{R})$ or to $\mathrm{SO}_Q(\mathbb{R})^\circ$. Now note

$$\mathrm{SO}_Q(\mathbb{Z}) = \{g \in \mathrm{SO}_Q(\mathbb{R}) \,:\, g.\mathfrak{sl}_2(\mathbb{Z}) \subset \mathfrak{sl}_2(\mathbb{Z})\}$$

and so in particular $\mathrm{PGL}_2(\mathbb{Z})$ can we viewed as a subset of $\mathrm{SO}_Q(\mathbb{Z})$. As $\mathrm{PGL}_2(\mathbb{R})$ can be viewed as a finite index subgroup of $\mathrm{SO}_Q(\mathbb{R})$ it follows that $\mathrm{PGL}_2(\mathbb{Z})$ is a sub-lattice of $\mathrm{SO}_Q(\mathbb{Z})$ and thus has finite index.

This shows that we can replace in the statement of 4.1 the group $\mathrm{SO}_{\mathrm{disc}}(\mathbb{Z})$ by the image of $\mathrm{SL}_2(\mathbb{Z})$ under this injection. This implies

$$|\mathrm{SL}_2(\mathbb{Z})\backslash\{(\mathbb{Z}^2, dX^2 + \ell XY + dY^2) \hookrightarrow (\mathbb{Z}^3, \mathrm{disc})\}| \ll_\varepsilon f \max(|d|, |\ell|)^\varepsilon. \quad (4.1)$$

5. Proof of Proposition 1.1

Denote by $\mathscr{F}$ the fundamental domain of X given by

$$\mathscr{F} = \{(z, v) \in \mathbb{H} \times S^1 \text{ such that } |\mathrm{Re}(z)| \leq \tfrac{1}{2} \text{ and } |z| \geq 1\}$$

and by $\mathscr{F}'$ a slight extension of $\mathscr{F}'$ defined as

$$\mathscr{F}' = \{(z, v) \in \mathbb{H} \times S^1 \text{ such that } |\mathrm{Re}(z)| \leq 1 \text{ and } |z| \geq \tfrac{1}{2}\}.$$

We progress with a few preliminary observations. Let $x_1, x_2 \in X_{\leq H}$ such that $d_X(x_1, x_2) < \delta$. Write $x_i = \Gamma g_i$ for $i = 1, 2$ and $g_i \in \mathrm{PGL}_2(\mathbb{R})$. In order to bound the coefficients of g_i, we always assume that the matrix g_i has determinant ± 1.

We choose g_1 such that $g_1 \in \mathscr{F}$ and $g_2 \in \mathscr{F}'$ such that $d_G(g_1, g_2) < \delta$. We claim that $||g_i|| \ll H$. We assume without loss of generality that $\det(g_1) = 1$. We use the NAK decomposition of $\mathrm{SL}_2(\mathbb{R})$ in order to write

$$g_1 = \begin{pmatrix} a & t \\ 0 & a^{-1} \end{pmatrix} k,$$

for $a, t \in \mathbb{R}$ with $a \neq 0$ and for $k \in \mathrm{SO}_2(\mathbb{R})$. As $g_1 \in \mathscr{F} \cap X_{\leq H}$ we conclude that $\mathrm{Re}(g_1.i) \leq 1$ and $\frac{1}{2} \leq \mathrm{Im}(g_1.i) \leq H^2$. We note that

$$g_1.i = a^2 i + at$$

and so $\frac{1}{2} \leq |a| \leq H$ and $|t| \leq \frac{1}{|a|} \leq 2$. Thus all the coefficients of g_1 are $\ll H$. As g_2 is close to g_1 we conclude that $||g_i|| \ll H$.

We now associate to g_i the primitive integral quadratic form

$$q_i(X, Y) = \sqrt{d}[g_i.q_0](X, Y) = a_i X^2 + b_i XY + c_i Y^2$$

with $d = b_i^2 - 4a_i c_i$ and $\gcd(a_i, b_i, c_i) = 1$. Towards the estimate of Proposition 1.1 the case where $q_1 = q_2$ will be easier, so we focus on the case $q_1 \neq q_2$. We want to count the number of such possible tuples g_1, g_2 such that $q_1 \neq q_2$. The Γ-equivalence class of (q_1, q_2) only depends on the point x_1, x_2 or more precisely by the δ-neighborhood around x_1 and x_2. By compactness of $\mathcal{G}_d$ the number of distinct such quadratic form is finite and we write

$$\Gamma(q_1^{(1)}, q_2^{(1)}), \ldots, \Gamma(q_1^{(k)}, q_2^{(k)})$$

for a complete list of such quadratic forms. Our first aim is to count k effectively.

Lemma 5.1. *In the above setting,*

$$k \ll_\varepsilon d^{1+2\varepsilon} H^4 \delta^2.$$

Proof. As $||g_i|| \ll H$, it follows that

$$\max(|a_i|, |b_i|, |c_i|) \ll d^{1/2}H^2$$

and as by assumption $g_2 = g_1 h$ for $d(h, \mathrm{id}) < \delta$, we conclude that $q_2 = \sqrt{d}g_1(h.q_0)$ with $||h.q_0 - q_0|| \ll \delta$ and so

$$\max(|a_1 - a_2|, |b_1 - b_2|, |c_1 - c_2|) \ll d^{1/2}H^2\delta. \tag{5.1}$$

We now define the quadratic form

$$q(X, Y) = \mathrm{disc}(X(a_1, b_1, c_1) + Y(a_2, b_2, c_2)) = dX^2 + \ell XY + dY^2$$

for $\ell \in \mathbb{Z}$. Thus the map

$$\iota : \mathbb{Z}^2 \to \mathbb{Z}^3, \quad \begin{pmatrix} X \\ Y \end{pmatrix} \mapsto \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

defines a representation of $dX^2 + \ell XY + dY^2$ by the ternary form disc.

Note that

$$|2d - \ell| = |q(1, -1)| = \mathrm{disc}(a_1 - a_2, b_1 - b_2, c_1 - c_2) \ll dH^4\delta^2$$

and so there is only a finite number of possible values for ℓ. Furthermore, assuming that $q_1 \neq q_2$ we show that $\ell \neq \pm 2d$. Indeed, if $\ell = \pm 2d$, then

$$\begin{aligned} d(a_2 \mp a_1)^2 &= da_2^2 \mp 2da_2a_1 + da_1^2 = q(a_2, -a_1) \\ &= \mathrm{disc}(a_2(a_1b_1c_1) - a_1(a_2, b_2, c_2)) = (a_2b_1 - a_1b_2)^2 \end{aligned}$$

which contradicts the assumption that d is not a perfect square. So by (4.1),

$$N_{\ell,d} := |\mathrm{SL}_2(\mathbb{Z})\backslash\{(\mathbb{Z}^2, dX^2 + \ell XY + dY^2) \hookrightarrow (\mathbb{Z}^3, \mathrm{disc})\}| \ll_\varepsilon f \max(|d|, |\ell|)^\varepsilon$$

and so $N_{\ell,d} \ll fd^\varepsilon$ as $d \geq 0$ and as by (5.1)

$$|\ell| \ll |\ell - 2d| + 2d \ll 2d + dH^4\delta^2 \ll d$$

as $d^{-1/4} \leq \delta \leq H^{-2}$.

If $\Gamma(q_1^{(i)}, q_2^{(i)})$ and $\Gamma(q_1^{(j)}, q_2^{(j)})$ are different then they define different embeddings up to $\mathrm{SL}_2(\mathbb{Z})$-equivalence, where we view $\mathrm{SL}_2(\mathbb{Z}) \hookrightarrow \mathrm{SO}_{\mathrm{disc}}(\mathbb{Z})$. Thus

$$\begin{aligned} k &\leq \sum_{\text{all possible } \ell} N_{\ell,d} \\ &\leq \sum_{f^2|d} \sum_{\text{all possible } \ell} fd^\varepsilon \\ &\ll \sum_{f^2|d} fd^\varepsilon \frac{dH^4\delta^2}{f^2} \\ &\ll \sum_{f^2|d} d^{1+\varepsilon}H^4\delta^2 \\ &\ll_\varepsilon d^{1+2\varepsilon}H^4\delta^2, \end{aligned}$$

where in the third line we used that ℓ has to satisfy $|2d-\ell| \ll dH^4\delta^2$ and $f^2|\ell$ so that $\frac{\gcd(\ell,\delta)}{f^2}$ is square-free. In the last line we used that the number of divisors of d can be bounded by $\ll_\varepsilon d^\varepsilon$ for any $\varepsilon > 0$. $\square$

Lemma 5.2. *Let $(x_1, x_2) = (\Gamma g_1, \Gamma g_2) \in (\mathcal{G}_d \cap X_{\leq H})^2$ be as above such that $d_X(x_1, x_2) < \delta$ and $q_1 \neq q_2$. Then there is some j so that $x_1 = \Gamma g_1^{(j)} a_t$ where $t \in I_j$ for I_j some interval of length $\ll \log(d)$.*

Proof. Choose j so that $(\sqrt{d}(g_1.q.0), \sqrt{d}(g_2.q.0)) = (q_1^{(j)}, q_2^{(j)})$. We note that $\mathcal{G}_d \subset X_{<d^{1/4}}$ and so using (5.1) there is a constant c so that

$$\max(|a_1-a_2|, |b_1-b_2|, |c_1-c_2|) \leq cd^{1/2}(d^{1/4})^2\delta \leq cd\delta.$$

Thus $d(g_1a_t, g_2a_t') \geq \frac{1}{2c}d^{-1}$. In particular, $d(g_1a_t, g_2A) \geq \frac{1}{2c}d^{-1}$, so

$$d(g_1a_t, g_2A) \gg d^{-1}. \tag{5.2}$$

We now want to show that the inequality $d_G(g_1^{(j)}a_t, g_2^{(j)}A) < 1$ can only hold for t in an interval of length $\ll \log(d)$. To see this denote by $h = (g_1^{(j)})^{-1}g_2$ and we write

$$h = \begin{pmatrix} h_1 & h_2 \\ h_3 & h_4 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$$

with $||h|| \ll \delta$. Using left invariance of the metric, (5.2) is equivalent to

$$d_G\left(\begin{pmatrix} \pm e^t & 0 \\ 0 & \pm e^{-t} \end{pmatrix}, \begin{pmatrix} \pm e^s h_1 & \pm e^{-s}h_2 \\ \pm e^s h_3 & \pm e^{-s}h_4 \end{pmatrix}\right) \gg d^{-1} \tag{5.3}$$

for all $s, t \in \mathbb{R}$. We first show that $|h_2|, |h_3| \gg d^{-1}$. For a contradiction assume without loss of generality that $|h_2| \leq cd^{-1}$ for some small c determined later, then we have that

$$|h_4| = \frac{|1+h_2h_3|}{|h_1|} \ll \frac{1+cd^{-1}\delta}{|h_1|}.$$

We can assume that h_1 and h_4 are positive. For $s=0$ in (5.3), we choose t so that $e^t = h_1$. Then we can choose c small enough so that

$$|e^{-t} - h^4| = \frac{|cd^{-1}\delta|}{|h_1|} \ll d^{-1}$$

and so we arrive at a contradiction to (5.3). Thus $|h_2|, |h_3| \gg d^{-1}$.

Assume that $s, t \in \mathbb{R}$ so that

$$d_G\left(\begin{pmatrix} \pm e^t & 0 \\ 0 & \pm e^{-t} \end{pmatrix}, \begin{pmatrix} \pm e^s h_1 & \pm e^{-s}h_2 \\ \pm e^s h_3 & \pm e^{-s}h_4 \end{pmatrix}\right) < 1 \tag{5.4}$$

Then we have that $|e^s h_2| \ll 1$ and $|e^{-s}h_3| \ll 1$ and so in particular $e^s \ll \frac{1}{|h_2|} \ll d$ and $e^{-s} \ll \frac{1}{|h_3|} \ll d$ and so we have that $|s| \ll \log(d)$. So we already see that s can only be inside an interval of length $\ll \log(d)$. So we have that $|\pm e^t \mp e^s h_1| \ll 1$ and so $|e^t| \ll 1 + |e^s h_1| \ll 1 + \log(d)\delta \ll \log(d)$

for H and so d large enough (and in particular δ small). The same again holds for e^{-t} and so we conclude $|t| \ll \log(d)$. $\square$

Proof. (of Proposition 1.1) Let $(x_1, x_2) = (\Gamma g_1, \Gamma g_2) \in (\mathcal{G}_d \cap X_{\leq H})^2$ be such that $d_X(x_1, x_2) < \delta$ for g_1, g_2 as above. Then as before we associate to x_i the quadratic form $q_i = \sqrt{d}(g_i.q_0)(X, Y)$. Recall that $\text{length}(\mathcal{G}_d^2) = d^{1+o(1)}$. So it suffices to show that

$$\text{length}(\{(x, y) \in (\mathcal{G}_d \cap X_{\leq H})^2 \,:\, d_X(x, y) < \delta\}) \ll_\varepsilon H^4 \delta^3 d^{1+\varepsilon}$$

If $q_1 = q_2$ then x_1 and x_2 lie on the same geodesic and so in this case we have that all these points can be described by

$$\{(x, xa_t) \in (\mathcal{G}_d \cap X_{\leq H})^2 \,:\, |t| \ll \delta\} \subset \{(x, xa_t) \in (\mathcal{G}_d)^2 \,:\, |t| \ll \delta\}.$$

Thus we have $\text{length}(\{(x, xa_t) \in (\mathcal{G}_d \cap X_{\leq H})^2 \,:\, |t| \ll \delta\}) \ll \delta \cdot \text{length}(\mathcal{G}_d) \ll_\varepsilon \delta d^{\frac{1}{2}+\varepsilon}$ so as $d^{-1/4} \leq \delta$,

$$(\mu_d \times \mu_d)(\{(x, xa_t) \in (\mathcal{G}_d \cap X_{\leq H})^2 \,:\, |t| \ll \delta\}) \ll_\varepsilon \delta d^{-1/2} d^\varepsilon \ll_\varepsilon \delta^3 d^\varepsilon.$$

In the case $q_1 \neq q_2$ Lemma 5.1 and Lemma 5.2 apply and so the length of the set

$$\{(x_1, x_2) \in (\mathcal{G}_d \cap X_{\leq H})^2 \,:\, d_X(x_1, x_2) < \delta \text{ and } q_1 \neq q_2\}$$

can be bounded by $\sum_{j=1}^k |I_j| \delta \ll \log(d) k\delta \ll_\varepsilon H^4 \delta^3 d^{1+2\varepsilon}$. $\square$

References

[ELMV12] M. Einsiedler, E. Lindenstrauss, P. Michel, and A. Venkatesh, *The distribution of closed geodesics on the modular surface, and Duke's theorem*, Enseign. Math. (2) **58** (2012), no. 3-4, 249–313. ↑1, 2, 4, 5

[EW11] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, 259, Springer-Verlag London, Ltd., London, 2011. ↑2

[Kog] C. Kogler, *Number Theory and Group Theory with a view towards Homogeneous Dynamics.* In preparation, Available at `http://www.constantinkogler.com/Files/SemesterpaperCK.pdf`. ↑3, 4

[Kog19] C. Kogler, *Ideal, Periodic Orbits and Duke's Theorem*, 2019. Available at `http://www.constantinkogler.com/Files/DukeSeminar1VersionPGL.pdf`. ↑1, 2