TOPOLOGICAL CONJUGACY OF LINEAR ENDOMORPHISMS OF THE 2-TORUS

ROY ADLER, CHARLES TRESSER, AND PATRICK A. WORFOLK

Abstract. We describe two complete sets of numerical invariants of topological conjugacy for linear endomorphisms of the two-dimensional torus, *i.e.*, continuous maps from the torus to itself which are covered by linear maps of the plane. The trace and determinant are part of both complete sets, and two candidates are proposed for a third (and last) invariant which, in both cases, can be understood from the topological point of view. One of our invariants is in fact the ideal class of the Latimer-MacDuffee-Taussky theory, reformulated in more elementary terms and interpreted as describing some topology. Merely, one has to look at how closed curves on the torus intersect their image under the endomorphism. Part of the intersection information (the intersection number counted with multiplicity) can be captured by a binary quadratic form associated to the map, so that we can use the classical theories initiated by Lagrange and Gauss. To go beyond the intersection number, and shortcut the classification theory for quadratic forms, we use the rotation number of Poincaré.

1. Introduction

The main aim of this paper is to approach the problem of topological conjugacy of linear endomorphisms of the 2-torus $\mathbf{T}^2 = \mathbf{R}^2/\mathbf{Z}^2$ from the topological point of view. The map $f : \mathbf{T}^2 \to \mathbf{T}^2$ is a *linear endomorphism* if it has a linear lift $F : \mathbf{R}^2 \to \mathbf{R}^2$ to the universal cover $\mathbf{R}^2$ of $\mathbf{T}^2$. In the rest of the paper, *endomorphism* means linear endomorphism of $\mathbf{T}^2$ and all lifts are assumed to be linear. Let M_F be the matrix representing F in some coordinate system; then M_F is a **Z**-*matrix*, *i.e.*, a matrix with integer entries. Let G be a lift of the endomorphism g and M_G the corresponding **Z**-matrix. Then it is well known that f is topologically conjugate to g if and only if the conjugacy can be chosen to be a linear automorphism (see [1] for the case of automorphisms; in general, topological conjugacy implies linear conjugacy of the maps induced on the first homology group, but these are the maps that we start with), which in turn is equivalent to saying that M_F and M_G are conjugate by an element of $Gl(2, \mathbf{Z})$. At the matrix level, the conjugacy problem was considered in [22, 31, 32] which in particular prove that the number of conjugacy classes within a *similarity class* is finite. Here *similar* matrices are matrices which are $Gl(2, \mathbf{Q})$-conjugate. The invariants of conjugacy within a similarity class in [22, 31, 32] are arithmetical (ideal classes) and the methods of proof depend on the theory of algebraic number fields, which allow the results to carry over to arbitrary dimension under fairly general assumptions. We propose for dimension two to exhibit complete invariants of conjugacy which can be understood from the topological point of view, and to get them by elementary methods. We describe how

Date: July 24, 1996.
1991 *Mathematics Subject Classification.* 58F35, 15A36, 11E16.

to solve the conjugacy problem, and present several normal forms for the conjugacy classes. In the particular important case of automorphisms, further normal forms derived from continued fractions are provided which allow one to easily construct interesting examples of similar but non-conjugate matrices. (See, for instance, Corollary 16 in Section 5 which tells when a matrix is conjugate to its transpose.)

Anyone familiar with the literature on binary quadratic forms will have a strong feeling of *"déjà vu"* when browsing through our Sections 3 to 5. The reason for that will become obvious in Section 6 where we reduce our conjugacy problem to the classification of quadratic forms, a theory mostly completed by Lagrange [19, 20] and Gauss [12], and then polished by several others. The quadratic form which we attach to a matrix is already known, at least for matrices in $Sl(2, \mathbf{Z})$. We learned this first from Etienne Ghys, but see, for instance, Katok [15].

In Section 7, we provide an alternate approach to the conjugacy problem which uses Poincaré's rotation number [27]. This is probably the most original part of the paper. We chose not to concentrate solely on this since using the material in Sections 3–5, Section 6 seems to lead to an introduction to the classical theory of quadratic forms with a topological flavor.

2. The conjugacy problem

The problem of determining conjugacy in $Gl(n, \mathbf{Z})$ of similar integral matrices with (the same) irreducible characteristic polynomial is identical with that of determining equivalence of ideals in the ring $\mathbf{Z}(\lambda)$ where λ is a root of the characteristic polynomial. Thus, because the number of ideal classes in this ring is known to be finite, so is the number of conjugacy classes of matrices satisfying the above hypothesis. This is the idea behind an important particular case of the Latimer-MacDuffee-Taussky theorem (for definitions see [6] which also contains an exposition of this theory by Olga Taussky; and for a treatment of the classical theory of quadratic forms, the reader can consult, *e.g.,* [5, 6, 25, 26] or the old masters).

The connection between conjugacy of matrices satisfying the above hypothesis and equivalence of ideals is a consequence of the following statements. On one hand each ideal in the ring $\mathbf{Z}(\lambda)$ has a $\mathbf{Z}$-basis. Such bases consist of n ring elements; and for the purposes of this discussion, let us call a vector whose components are the set of elements of a $\mathbf{Z}$-basis, a $\mathbf{Z}$-*basis vector*, and mean by scalar multiplication that by a ring element. An element of $Gl(n, \mathbf{Z})$ transforms one $\mathbf{Z}$-basis vector of an ideal to another of the same ideal; and conversely, for any two $\mathbf{Z}$-basis vectors of an ideal there exists an element of $Gl(n, \mathbf{Z})$ transforming one to the other. Two ideals are equivalent if some scalar multiple of any $\mathbf{Z}$-basis vector of the first ideal generates the same ideal as some scalar multiple of any $\mathbf{Z}$-basis vector of the second. These two new vectors are themselves $\mathbf{Z}$-basis vectors of the same ideal and thus one is transformed to the other by an element of $Gl(n, \mathbf{Z})$. On the other hand, any $n \times n$ integral matrix satisfying the aforementioned hypothesis has an eigenvector associated with the eigenvalue λ which is the $\mathbf{Z}$-basis vector of an ideal and all equivalent ideals are generated by scalar multiples of this eigenvector. Conversely, λ times a $\mathbf{Z}$-basis vector of an ideal is the image of that vector by some $n \times n$ integer matrix satisfying the hypothesis. Finally, an n-dimensional version of our forthcoming Lemma 1 says that two similar $n \times n$ integral matrices A and B satisfying the above hypothesis are conjugate if and only if some element in $Gl(n, \mathbf{Z})$ maps an eigenvector of A associated with λ to one of B.

Even though the connection between conjugacy and ideal equivalence provides some information, what we want is a practical way to decide if two similar **Z**-matrices are conjugate. To this end, we restrict our attention to 2×2 matrices (although of course, we would love to get rid of this restriction).

In full detail in dimension 2, the conjugacy problem is the following. Given two **Z**-matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},\ B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}\ ,$$

find a matrix $C = \begin{pmatrix} x & y \\ u & v \end{pmatrix} \in Gl(2, \mathbf{Z})$ such that $CA = BC$. This is equivalent to solving the linear set of equations given by

$$\begin{pmatrix} ax + cy & bx + dy \\ au + cv & bu + dv \end{pmatrix} = \begin{pmatrix} ex + fu & ey + fv \\ gx + hu & gy + hv \end{pmatrix}\ ,$$

subject to the nonlinear condition

$$xv - uy = \pm 1\ .$$

We can eliminate u, v and find that x, y must satisfy the following:

$$\begin{aligned} bx^2 + (d-a)xy - cy^2 &= \pm f \\ (a-e)x + cy &\equiv 0 \mod f \\ bx + (d-e)y &\equiv 0 \mod f\ . \end{aligned} \tag{1}$$

Alternatively, eliminating x, y leaves:

$$\begin{aligned} bu^2 + (d-a)uv - cv^2 &= \mp g \\ bu + (d-h)v &\equiv 0 \mod g \\ (a-h)u + cv &\equiv 0 \mod g\ . \end{aligned} \tag{2}$$

It is easy to check that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\ .$$

It is also easy to check that the two *companion matrix* forms in a similarity class are conjugate — namely,

$$\begin{pmatrix} -b & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix} = \begin{pmatrix} 0 & a \\ 1 & b \end{pmatrix}\begin{pmatrix} -b & 1 \\ 1 & 0 \end{pmatrix}\ .$$

We now consider an example given by Boyle & Handelman [3]. The matrix $\begin{pmatrix} -1 & 3 \\ 3 & 2 \end{pmatrix}$ is not conjugate to its companion matrix $\begin{pmatrix} 0 & 1 \\ 11 & 1 \end{pmatrix}$. This is evident since it leads to solving the integer quadratic form from (1),

$$3x^2 + 3xy - 3y^2 = \pm 1\ ,$$

which clearly has no integer solutions since 3 does not divide 1.

The matrix $\begin{pmatrix} 3 & 10 \\ 5 & 17 \end{pmatrix}$ is similar but not conjugate to its inverse. This is no longer a quick check using the above equations. In this paper, we discuss the way to verify results of this type.

3. Finiteness of the number of classes

Given a 2×2 **Z**-matrix, we denote the trace by T, the determinant by D, and the discriminant by $\Delta \, (\equiv T^2 - 4D)$. We shall restrict our attention to **Z**-matrices with irreducible characteristic polynomial (the non-trivial case) and denote their set by $\mathcal{M}$:

$$\mathcal{M} \equiv \{2 \times 2 \ \mathbf{Z}\text{-matrices} | \, \Delta \neq n^2 \, , n \in \mathbf{Z}\} \, .$$

We shall frequently discriminate based on the sign of the discriminant, so let

$$\begin{aligned} \mathcal{M}^+ &= \{M \in \mathcal{M} | \Delta > 0\} \, , \\ \mathcal{M}^- &= \{M \in \mathcal{M} | \Delta < 0\} \, . \end{aligned}$$

If $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}$, then the eigenvalues are given by

$$\lambda_\pm = \frac{T \pm \sqrt{\Delta}}{2} = \frac{a + d \pm \sqrt{(a+d)^2 - 4(ad - bc)}}{2} \, ,$$

with corresponding (real or complex) eigenvectors

$$\begin{pmatrix} b \\ \lambda_\pm - a \end{pmatrix} \, .$$

The pair of eigenvalues, or equivalently the trace and the determinant, are invariants of the similarity class. Consequently, the discriminant is also a similarity class invariant. In the case of irreducible characteristic polynomial which we study here, the trace and determinant form a complete set of similarity invariants.

Let us denote conjugacy (by elements of $Gl(2, \mathbf{Z})$) of similar matrices M, N by $M \sim N$. The restriction to conjugacy by elements of $Sl(2, \mathbf{Z})$ is called *proper* conjugacy. If two similar matrices are conjugate, then eigenvectors corresponding to the same eigenvalue are mapped to each other by the conjugacy. The converse of this for elements of $\mathcal{M}$ is stated in the following lemma.

Lemma 1. *Let $M, N \in \mathcal{M}$ be similar. If an element A of $Gl(2, \mathbf{Z})$ maps one eigenvector of M to the corresponding eigenvector of N, then $M = A^{-1}NA$.*

Proof. Let v be an eigenvector of M corresponding to an eigenvalue μ and assume that $A \in Gl(2, \mathbf{Z})$ is given such that Av is an eigenvector of N corresponding to the eigenvalue μ. Let $\bar{\mu}, \bar{v}$ be the (quadratic) conjugates of μ, v, then $\bar{v}$ and $A\bar{v}$ are eigenvectors of M and N, respectively, corresponding to eigenvalue $\bar{\mu}$. From the irreducibility of the characteristic polynomial of $M \in \mathcal{M}$, the eigenvectors v and $\bar{v}$ are linearly independent. We then get that $M - A^{-1}NA$ is identically zero by applying it to v and $\bar{v}$. $\square$

We define the *principal slope* ω_M of the element M of $\mathcal{M}$ to be the slope of the eigenvector corresponding to λ_+:

$$\omega_M = \frac{\lambda_+ - a}{b} = \frac{d - a + \sqrt{\Delta}}{2b} \, . \tag{3}$$

Note that ω_M is either complex or a quadratic surd (an irrational root of a quadratic equation with integer coefficients). The slope of the other eigenvector is denoted by $\overline{\omega_M}$ and is the (complex or quadratic) conjugate of ω_M. If several matrices are involved in a discussion, we use subscripts on the correponding (principal) slopes.

Recall that two complex numbers ω, ω' are *equivalent*, which we denote by $\omega \sim \omega'$, if they belong to the same orbit of the group $PGl(2,\mathbf{Z})$ acting on $\mathbf{C}$ by fractional linear transformations. If $\Gamma \in PGl(2,\mathbf{Z})$ and

$$\Gamma : \omega \mapsto \frac{\alpha\omega+\beta}{\gamma\omega+\delta},\ \alpha\delta-\beta\gamma=\pm 1,\ \alpha,\beta,\gamma,\delta \in \mathbf{Z}$$

then we associate Γ with the matrix $M_\Gamma \equiv \begin{pmatrix} \delta & \gamma \\ \beta & \alpha \end{pmatrix}$. Lemma 1 can now be reformulated:

Lemma 2. *Let M, $N \in \mathcal{M}$ be similar. Then $M \sim N$ if and only if $\omega_M \sim \omega_N$. In particular, $\omega_N = \Gamma(\omega_M)$ if and only if $N = M_\Gamma M M_\Gamma^{-1}$.*

To each element M of $\mathcal{M}$ we associate a set S_M defined as follows. If $\Delta > 0$ then $\omega, \overline{\omega}$ are real and we define the set S_M to be the circle in the complex plane through the diametrically opposite points ω and $\overline{\omega}$. If $\Delta < 0$ then $\omega, \overline{\omega}$ are complex and we define the set S_M to be simply the pair of points $\{\omega, \overline{\omega}\}$. For any fundamental region R for the action of $PGl(2,\mathbf{Z})$ on $\mathbf{C}\setminus\mathbf{R}$, we say that M is *reduced* if S_M intersects R. (Recall that $PGl(2,\mathbf{Z})$ acts transitively on the set difference $\mathbf{C}\setminus\mathbf{R}$ which correpsonds to the better known transitive action of $PSl(2,\mathbf{Z})$ on the upper half plane.) After noting that every element of M is conjugate to a reduced matrix, it is then clear that the number of reduced matrices in a similarity class gives an upper bound on the number of conjugacy classes. In the sequel, we suppose for definiteness that R is the *principal fundamental region* R_0 defined by

$$R_0 = \{z \in \mathbf{C} \mid -\frac{1}{2} \le \Re z \le 0,\ \Im z > 0,\ |z| \ge 1,\ \text{or } 0 < \Re z < \frac{1}{2},\ \Im z > 0,\ |z| > 1\}\,.$$

Theorem 3. *Within a given similarity class in $\mathcal{M}$, the number of conjugacy classes is finite.*

Proof. In the complex case ($\Delta < 0$), there are only two reduced matrices. One has $\omega \in R_0$ and the other has $\overline{\omega} \in R_0$.

In the real case ($\Delta > 0, \Delta \ne n^2$), given a matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}$ with real eigenvalues, the circle S_M is the boundary of the disk $|z - \frac{d-a}{2b}| \le |\frac{\sqrt{\Delta}}{2b}|$. Now the circle S_M intersects R_0 only if the disk it bounds contains at least one of the corners $\pm\frac{1}{2} + i\frac{\sqrt{3}}{2}$ of R_0. Calculate

$$|\pm\frac{1}{2} + i\frac{\sqrt{3}}{2} - \frac{d-a}{2b}| \le |\frac{\sqrt{\Delta}}{2b}|\,,$$

or

$$((T-2a)\pm b)^2 + 3b^2 \le \Delta\,.$$

So, in a given similarity class, there are only a finite number of permissible values for a and b for which a matrix is reduced. A matrix in a similarity class is determined by a and b, so there are only finitely many reduced matrices in a similarity class. $\square$

Remarks.

1. This proof is only a rewording of proofs given by Dedekind [9] in the complex case and by Klein [17] in the real case for the classification of binary quadratic forms, following work by Smith [29] (see also the work by Humbert [14] who gives credit to Hermite for the real case, and the accounts given in Chapter IV of Mathews [25] and Chapter I of LeVeque [24]).

2. Notice that in these classical works, it is the counterpart of our proper conjugacy which is considered. Consequently, instead of $PGl(2,\mathbf{Z})$ acting on $\mathbf{C}$, it is $PSl(2,\mathbf{Z})$ acting on the upper half-space which has to be considered (with the same fundamental region R_0). In the complex case, one chooses from ω and $\bar{\omega}$ the one with positive imaginary part, yielding a single reduced form per class.

3. The number of proper equivalence classes in a given similarity class is at most twice the number of equivalence classes, hence finite.

4. The intersection of the set S_M with the upper half-plane for $\Delta > 0$ is a semicircle which is a geodesic in hyperbolic geometry. Because there are only finitely many reduced forms, the projection of this curve into the fundamental domain yields a closed curve. An example is presented in Figure 1. We have selected $\Delta = 40$ and this results in two geodesics, i.e., a similarity class of matrices in $Gl(2,\mathbf{Z})$ with discriminant $\Delta = 40$ contains exactly two conjugacy classes. Similar closed curves were already discussed by Smith [29], but with a different fundamental region.

4. Solution to the conjugacy problem

The result in Lemma 2 demonstrating that equivalence of principal slopes is equivalent to conjugacy of similar matrices can be used to solve the conjugacy problem. We will consider the cases of positive and negative discriminant separately.

4.1. Negative discriminant.

When $\Delta < 0$, ω_M and ω_N are in $\mathbf{C}\backslash\mathbf{R}$ and there are unique elements Γ_M and Γ_N in $PGl(2,\mathbf{Z})$ such that $\Gamma_M(\omega_M)$ and $\Gamma_N(\omega_N)$ belong to the principal fundamental region R_0 of $PGl(2,\mathbf{Z})$. Thus we have

Theorem 4. *Let $M,\ N \in \mathcal{M}^-$ be similar. Then*

$$M \sim N \iff \Gamma_M(\omega_M) = \Gamma_N(\omega_N)\,.$$

The computation of Γ_M is well documented (modulo the second remark in the previous section, see, *e.g.*, [25]), so we recall it only briefly. Essentially, we map ω_M into the upper half plane and then translate horizontally into the vertical strip of width one centered about the complex axis and containing the fundamental region R_0. If the image of ω_M is not in the fundamental domain because its magnitude is less than one, then we take the negative reciprocal which will increase the imaginary part. We then repeat the translation and reciprocal processes. One can check that eventually the point will be brought into R_0. The following algorithm implements this process.

Algorithm 5. Computation of Γ_M.

1. If $\Im(\omega_M) < 0$, let Γ be the map $z \mapsto \frac{1}{z}$, otherwise let $\Gamma = \mathrm{Id}$. Thus, $\Gamma(\omega_M)$ is in the upper half plane.
2. Let Γ_1 be the translation by an integer $z \mapsto z + n$ that maps $\Gamma(\omega_M)$ into the vertical strip $S:\ -\frac{1}{2} \leq \Re(z) < \frac{1}{2}\,,\ \Im(z) > 0$. Let $\Gamma = \Gamma_1 \circ \Gamma$.
3. If $\Gamma(\omega_M) \in R_0$, then let $\Gamma_M = \Gamma$ and we are done. Otherwise let Γ_2 be the map $z \mapsto \frac{-1}{z}$, let $\Gamma = \Gamma_2 \circ \Gamma$, and return to step 2.
   Remark that the map Γ_2 increases the imaginary part of any z in $S \setminus R_0$.

The finiteness of the process is guaranteed by the finiteness of the number of fundamental regions in S above any horizontal line in the upper half plane and the result stated at the end of step 3.

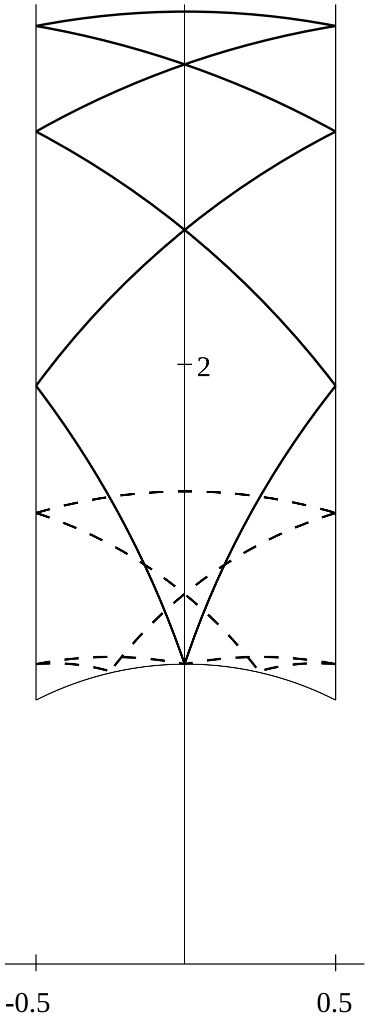

FIGURE 1. The two periodic hyperbolic geodesics corresponding to $\Delta = 40$.

4.2. **Positive non-square discriminant.** For matrices in $\mathcal{M}^+$ we use continued fractions in order to determine equivalence of principal slopes. We write uniquely the infinite continued fraction of an irrational number as $[a_0\ a_1\ a_2\ \ldots]$ with $a_i \in \mathbf{Z}$ for all i and $a_i > 0$ for all $i > 0$. By a theorem of Serret [28] (see also [13, Theorem 175], two real numbers are equivalent if and only if their continued fractions have the same tail. Lagrange [18] (see also [13, Theorem 177]) proved that quadratic surds have periodic continued fractions which we write as $[a_0\ a_1\ \ldots\ a_n\ \overline{b_1\ \ldots\ b_m}]$, where the overbar means an infinite repetition of b_1 through b_m. We define the *period* of a periodic continued fraction to be the minimal m such that the fraction may be written in this form. The length of the aperiodic part is $n+1$. The periodic tail of this continued fraction is the purely periodic continued fraction $[\overline{b_1\ \ldots\ b_m}]$, and if the period is m then the periodic part is the ordered set $\{b_1, \ldots, b_m\}$, which is well defined only up to cyclic permutation.

Remark. For continued fractions, one uses "periodic" and "purely periodic" where in dynamics one would use "pre-periodic" and "periodic" respectively; we use here the former convention. For a general reference on continued fractions see, *e.g.*, Davenport [8] or Hardy & Wright [13].

When Δ is positive and not a perfect square, the principal slopes ω_M, ω_N of M and N are quadratic surds, thus the continued fraction results allow us to easily prove

Theorem 6. *Let $M, N \in \mathcal{M}^+$ be similar. Then $M \sim N$ if and only if the periodic parts of the continued fraction of ω_M and ω_N are the same (up to a cyclic permutation).*

Remark. For a similar result for matrices in $Sl(2, \mathbf{Z})$, see Katok [15].

4.3. **Proper conjugacy.** If we want to consider only proper conjugacy, then the previous results from 4.1 and 4.2 have to be modified as follows:

If $\Delta < 0$, then conjugate matrices M and N are properly conjugate if and only if the imaginary parts of ω_M and ω_N have the same sign. Note that the imaginary parts of ω_M and ω_N will have the same sign if and only if the matrix elements $M_{1,2}$ and $N_{1,2}$ have the same sign.

If $\Delta > 0$ and $\Delta \neq n^2$, then a simple modification of the argument of Serret [28] (see also Lejeune-Dirichlet [23, p. 163] or Smith [30, p. 186]) yields:

Theorem 7. *Conjugate matrices $M, N \in \mathcal{M}^+$ are properly conjugate if and only if either the continued fractions of ω_M and ω_N have odd period or the continued fractions have even period and the difference in lengths of the aperiodic parts is even (where the lengths of the aperiodic parts must be counted so that the corresponding periodic parts are identical).*

5. Normal forms

The reduced matrices discussed in Section 3 are reasonable normal forms in the case of negative discriminant, but as in Section 4, we will require that the principal slope be in R_0. In the case of positive non-square discriminant, it is more convenient to choose matrices as being in normal form if the principal root (or its conjugate) has a purely periodic continued fraction: several choices will be proposed, each of which have different virtues.

5.1. **Normal forms for negative discriminant.** If $\omega_M \in R_0$, we have:

$$\begin{array}{l} \Re(\omega_M)^2 \leq \frac{1}{4} \text{ and } (\Re\omega_M)^2 + (\Im\omega_M)^2 \geq 1 \text{ if } \Re(\omega_M) \leq 0, \\ \Re(\omega_M)^2 < \frac{1}{4} \text{ and } (\Re\omega_M)^2 + (\Im\omega_M)^2 > 1 \text{ if } \Re(\omega_M) > 0\,. \end{array} \tag{4}$$

The corresponding conditions on the entries in the matrix M are given in the following lemma.

Lemma 8. *A matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}^-$ has its principal slope ω_M in the principal fundamental region R_0 if and only if*

$$-A < B \leq A < C \quad \text{or} \quad 0 \leq B \leq A = C\,,$$

where

$$A = b,\ B = a - d,\ C = -c\,.$$

There is exactly one such matrix in each conjugacy class.

Proof. One can check that the conditions in the theorem are equivalent to the conditions in (4) with ω_M defined by (3). Uniqueness is a result of R_0 being a fundamental domain for the action of $PGl(2,\mathbf{Z})$ on $\mathbf{C}\setminus\mathbf{R}$. $\square$

We thus define the *normal form* for a conjugacy class with $\Delta < 0$ to be the unique element with principal slope contained in R_0.

To conjugate a matrix $N \in \mathcal{M}^-$ to its normal form N_0, one uses Algorithm 5 to produce the unique mapping Γ_N and then conjugates N with the matrix M_{Γ_N} as specified in Lemma 2. In the case of proper conjugacy, nothing changes if $\Im(\omega_N) > 0$, but if $\Im(\omega_N) < 0$, one looks for the element $\overline{\Gamma_N} \in PSl(2,\mathbf{Z})$ which sends the conjugate $\overline{\omega_N}$ to R_0.

One is often interested in the special case of toral automorphisms which results in studying the conjugacy classes of $Gl(2,\mathbf{Z})$. When we consider elements in $Gl(2,\mathbf{Z})$ with negative discriminant there are only three similarity classes which (specifed by the determinant D and trace T) are:

$$(D,T) \in \{(1,0),(1,1),(1,-1)\}\ .$$

One can now study which elements in these similarity classes are reduced by using the criteria in Lemma 8. We leave it to the reader to verify that there is exactly one reduced matrix in each of the similarity classes and, thus, exactly one conjugacy class in each similarity class. We state this result in

Theorem 9. *Let $M \in \mathcal{M}^- \cap Gl(2,\mathbf{Z})$. The principal slope ω_M is in the principal fundamental region R_0 if and only if M is one of the following matrices:*

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

5.2. **Normal forms for positive non-square discriminant.** In the case of positive non-square discriminant, we present three different choices for normal forms, and in each case there is not in general a unique representative for each conjugacy class.

5.2.1. *Purely periodic principal slope.* Our first candidates for *normal forms* for a conjugacy class with $\Delta > 0$, $\Delta \neq n^2$, are the matrices whose principal slopes have purely periodic continued fractions. A conjugacy to one of these normal forms can be constructed from a sequence of conjugacies which are chosen so that they simply peel off, one term at a time, the aperiodic part of the continued fraction expansion of the principal slope. These conjugacies are explicitly given later in Lemma 13.

In the case of proper conjugacy, because the conjugacy which strips one term in the continued fraction is orientation reversing, the peeling off process must remove an even number of terms of the continued fraction. Thus, if the length of the period is even, only half of the purely periodic continued fractions equivalent to ω_M can be reached by proper conjugacy.

The following theorem by Galois [11] (see also [8, p. 99]) gives conditions on a quadratic surd for it to have a purely periodic continued fraction expansion.

Theorem 10 (Galois [11])**.** *Let x be a quadratic surd and $\overline{x}$ its conjugate.Then x has a purely periodic continued fraction if and only if*

$$-1 < \overline{x} < 0 \text{ and } x > 1\,.$$

Furthermore, the continued fraction of $-1/\overline{x}$ is purely periodic and has the same terms in its period as the continued fraction of x, but in the reverse order.

Remark. More generally if x is a quadratic surd and $\overline{x}$ is its conjugate, the periodic parts of the continued fractions of x and $\overline{x}$ have the same terms but in reverse order. This follows at once from Theorem 10 and the fact that if $x \sim x'$ with $x' = (ax+b)/(cx+d)$ then $-1/x \sim -1/x'$ since $-1/x' = (-d/x-c)/(b/x+a)$.

As a corollary to Galois's Theorem, we give explicit conditions on the elements in a matrix $M \in \mathcal{M}^+$ so that the principal slope of M will have a purely periodic continued fraction expansion.

Corollary 11. *A matrix* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}^+$ *has a principal slope whose continued fraction is purely periodic if and only if*

$$\begin{aligned} &b > 0\,, \\ &c > 0\,, \\ &d > a + |b-c|\,. \end{aligned}$$

Proof. A direct calculation using ω_M shows that the above conditions are necessary and sufficient to show that $1 < \omega_M$ and $-1 < \overline{\omega}_M < 0$. By Theorem 10, ω_M then has a purely periodic continued fraction. □

5.2.2. *Purely periodic principal slope or conjugate.* Another choice (for conjugacy or proper conjugacy) consists in considering M as a normal form if and only if ω_M or $\overline{\omega}_M$ is purely periodic. This gives twice as many normal forms for conjugacy or proper conjugacy (hence an even number of normal forms in both cases).

If we interpret our matrices as mappings on $\mathbf{R}^2$, then we can consider the associated phase portraits. It follows from Galois's theorem that this normal form corresponds to having any one of the phase portraits represented in Figure 2, where, as usual, more arrows on an invariant subspace corresponding to an eigenvector means a bigger absolute value for the corresponding eigenvalue. In the top row, the slope corresponding to the eigenvalue with larger magnitude has been chosen to be purely periodic, while in the bottom row the slope of the smaller (in magnitude) eigenvalue is purely periodic. If the eigenvalues have the same magnitude ($\mathrm{tr}(M) = 0$), then the phase portraits I and I' are the same (and the saddle-type phase portraits do not occur). A phase portrait with two contracting directions is not possible since the matrices we consider do not have determinant with magnitude less than one. Reinterpreting Figure 2 in terms of our first kind of normal forms, the eigendirection in the first quadrant is the one corresponding to ω_M.

5.2.3. *On a Theorem by R. F. Williams.* For our third definition of normal forms we select phase portraits I and II of Figure 2. This means that when $\mathrm{tr}(M) > 0$ then we require that ω_M has a purely periodic continued fraction, but when $\mathrm{tr}(M) < 0$ then we require that $\overline{\omega}_M$ has a purely periodic continued fraction. (When $\mathrm{tr}(M) = 0$, then either slope may be purely periodic.) The benefit of this third choice is that we get in many cases normal forms which are either nonpositive or nonnegative.

Theorem 12. *Let* $M \in \mathcal{M}^+$ *which either preserves orientation or has a saddle-type phase portrait. If* $tr(M) > 0$ *and* ω_M *has a purely periodic continued fraction, then* M *is a non-negative matrix. If* $tr(M) < 0$ *and* $\overline{\omega}_M$ *has a purely periodic continued fraction, then* M *is a non-positive matrix.*

Proof. If $\mathrm{tr}(M) > 0$ and ω_M has a purely periodic continued fraction or if $\mathrm{tr}(M) < 0$ and $\overline{\omega}_M$ has a purely periodic continued fraction, then M has a phase portrait corresponding to figures I or II. We will now show that if M has a phase portrait

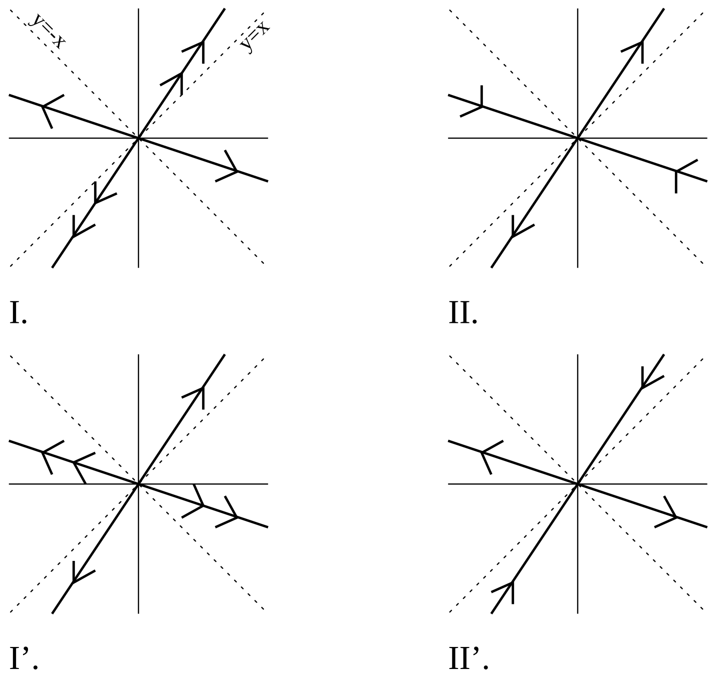

FIGURE 2. Phase portraits of normal forms.

corresponding to I or II, then its elements do not have opposite signs. The conditions on the signs of the elements of M are equivalent to the requirement that the first quadrant is either mapped into itself or into the third quadrant. It is easily verified that if M is orientation preserving then the maps producing phase portraits I and II satisfy this requirement. Thus the only nontrivial case is when M has a saddle-type phase portrait and is orientation reversing.

We first examine this case when the trace of M is positive, so that we assume $-1 < \lambda_- < 0$, $\lambda_+ > 1$, and we have to look closer at case II in Figure 2.

Assume the theorem is false. We let P be the point $(1, 0)$ and choose Q on the unstable direction such that PQ is parallel to the stable direction. We let R be the point of the unstable direction such that $M(P)R$ is horizontal. Then in Figure 3a, the triangles OPQ and $M(P)RM(Q)$ are similar.

Now $|M(P)M(Q)| = |\lambda_-| \cdot |QP| < |QP|$, so that $|M(P)R| < OP = 1$. Since M is a $\mathbf{Z}$-matrix, $M(P)$ is in $\mathbf{Z}^2$, so that the distance from $M(P)$ to R is at least 1 if $M(P)$ is out of the first quadrant. Hence $M(P)$ does belong to the first quadrant, to the contrary of what is represented in Figure 3a.

On the other hand, with T standing for the point $(0, 1)$, it is plain that $M(T)$ is also in the first quadrant since the slope of the unstable direction is greater than 1. It follows that the coefficients of the matrix M are non-negative.

If we assume that the trace is negative, we replace Figure 3a by Figure 3b. Then we conclude similarly that the image of the basis is in the third quadrant, so that the coefficients are non-positive. $\square$

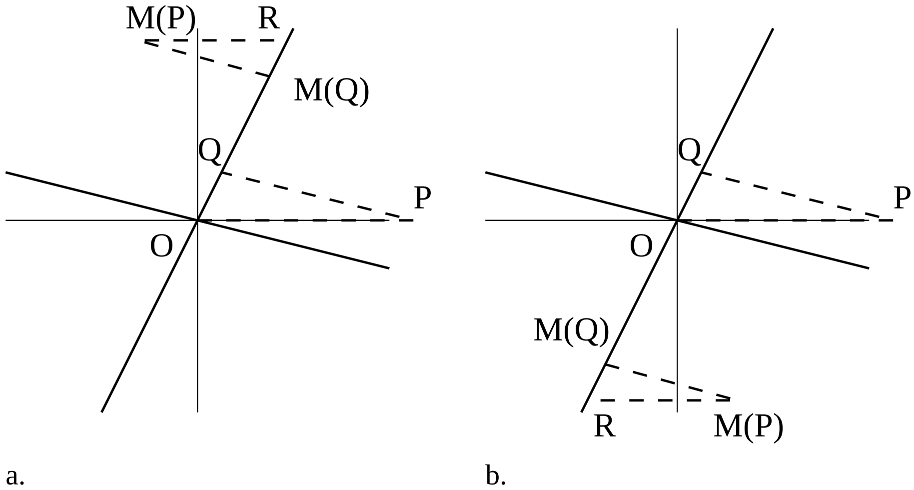

FIGURE 3. Impossible situation for case II with: a. positive trace; b. negative trace.

A direct result of this theorem is of course that if $M \in \mathcal{M}^+$ and M is either orientation preserving or has a saddle-type phase portrait, then M may be conjugated to a matrix whose elements do not have opposite signs.

Remarks.

1. Williams [33] proved that hyperbolic elements of $Gl(2,\mathbf{Z})$ can be conjugated so that the resulting matrix is either nonpositive or nonnegative. These are just the elements of $\mathcal{M}^+$ with determinant $D = \pm 1$. Since these all have saddle-type phase portraits, they are all conjugate to matrices which have phase portraits of type II, and hence are also included in our result.

2. The matrix $\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} \in \mathcal{M}^+$ is orientation reversing and does not have a saddle-type phase portrait. It has positive trace and the continued fraction of its principal slope is purely periodic (by Corollary 11), but it is not a non-negative matrix.

3. The matrix $\begin{pmatrix} -1 & 3 \\ 3 & 2 \end{pmatrix} \in \mathcal{M}^+$ from Section 2 is not conjugate to any non-negative or non-positive matrix. The only non-negative matrices it is similar to are its companion matrices, but it is not conjugate to these.

4. For $M \in \mathcal{M}^+$, preserving orientation or with saddle-type phase portrait, one cannot have $\mathrm{tr}(M) = 0$. Notice that there are examples with trace zero not conjugate to a nonpositive or nonnegative matrix; for instance $\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$ is not conjugate to $\begin{pmatrix} 0 & 5 \\ 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & -5 \\ -1 & 0 \end{pmatrix}$ or their transposes as the method in Section 2 easily shows.

5.2.4. *The special case of automorphisms.* The reduction of any matrix in $\mathcal{M}^+$ is mirrored by the reduction of a matrix in $Gl(2,\mathbf{Z})$. This follows from the fact that any quadratic surd can be found as the principal slope of a matrix in $Gl(2,\mathbf{Z})$.

We will also demonstrate how to construct all matrices in $Gl(2,\mathbf{Z})$ with a specified periodic continued fraction for their principal slopes. To prove these results we first need two lemmas.

Let

$$T_a = \begin{pmatrix} 0 & 1 \\ 1 & a \end{pmatrix} .$$

Lemma 13. *Let $A \in \mathcal{M}^+$ and $\omega_A = [a_0\ a_1\ a_2\ \ldots]$. Then $\omega_{T_{a_0}^{-1} A T_{a_0}} = [a_1\ a_2 \ldots]$. If also $a_0 > 0$, then $\omega_{T_b A T_b^{-1}} = [b\ a_0\ a_1\ a_2\ \ldots]$.*

Recall that when we write a continued fraction as $[a_0\ a_1\ a_2\ \ldots]$ then each a_i, for $i > 0$, is a positive integer, but a_0 may be any integer.

Proof. The results follow from straightforward applications of Lemma 2. $\square$

Lemma 14. *Let $A = (T_{b_1} \cdots T_{b_m})^N$ with $b_i > 0$ and $N \geq 1$ and maximal. Then $\omega_A = [\overline{b_1 \ldots b_m}]$.*

Proof. We leave it to the reader to show that nonnegative matrices of this type map the positive quadrant of $\mathbf{R}^2$ into itself and have a unique eigenvector in the positive quadrant which has slope greater than one.

Assume that $w_A = [a_0\ a_1\ a_2\ \ldots]$. Since $w_A > 1$, $a_0 > 0$ and we may apply the result in Lemma 13, to show

$$\omega_{T_{b_1} \cdots T_{b_m} A T_{b_m}^{-1} \cdots T_{b_1}^{-1}} = [b_1\ \ldots\ b_m\ a_0\ a_1\ a_2\ \ldots] .$$

But the above matrix equals A and by the uniqueness of continued fraction expansions, we see that ω_A has the requisite continued fraction. $\square$

Theorem 15. *Let $a_i > 0$ for $i > 0$, and let all $b_i > 0$. If*

$$A = T_{a_0} \cdots T_{a_n} (T_{b_1} \cdots T_{b_m})^N T_{a_n}^{-1} \cdots T_{a_0}^{-1}$$

for any positive integer N, then

$$\omega_A = [a_0\ \ldots\ a_n\ \overline{b_1\ \ldots\ b_m}] .$$

If

$$A = -T_{a_0} \cdots T_{a_n} (T_{b_1} \cdots T_{b_m})^N T_{a_n}^{-1} \cdots T_{a_0}^{-1}$$

for any positive integer N, then

$$\overline{\omega_A} = [a_0\ \ldots\ a_n\ \overline{b_1\ \ldots\ b_m}] .$$

Conversely, let $A \in \mathcal{M}^+ \cap Gl(2,\mathbf{Z})$. If $tr(A) > 0$ and

$$\omega_A = [a_0\ \ldots\ a_n\ \overline{b_1\ \ldots\ b_m}]$$

where m is the period, then

$$A = T_{a_0} \cdots T_{a_n} (T_{b_1} \cdots T_{b_m})^N T_{a_n}^{-1} \cdots T_{a_0}^{-1}$$

for some positive integer N; if $tr(A) < 0$ and

$$\overline{\omega_A} = [a_0\ \ldots\ a_n\ \overline{b_1\ \ldots\ b_m}]$$

where m is the period, then

$$A = -T_{a_0} \cdots T_{a_n} (T_{b_1} \cdots T_{b_m})^N T_{a_n}^{-1} \cdots T_{a_0}^{-1}$$

for some positive integer N.

Proof. The first two parts of the theorem follow directly from Lemmas 13 and 14 and the fact that $\overline{\omega_A} = \omega_{-A}$.

To prove the converse parts, we first note that A has a saddle type phase portrait and prove the result for the case when $\mathrm{tr}(A) > 0$. The other case follows from working with $-A$. So assume that $\mathrm{tr}(A) > 0$ which means that $|\lambda_+| > 1$. Define $\tilde{A} = T_{a_n}^{-1} \dots T_{a_0}^{-1} A T_{a_0} \dots T_{a_n}$. By Lemma 13, $\omega_{\tilde{A}} = [\overline{b_1 \ \dots \ b_m}]$. If $\tilde{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then because the continued fraction of its principal slope is purely periodic, we may apply Theorem 11 to conclude that $b, c, d - a > 0$. The phase portrait of $\tilde{A}$ is that of case II in Figure 2 due to the fact that the slope is purely periodic and $|\lambda_+| > 1$. From the proof of Theorem 12, we know that the signs of a, b, c, d cannot be opposite so that in addition to $b, c > 0$ we find that $d > a \geq 0$. We will show that $\tilde{A}$ is a product of matrices $T_\alpha = \begin{pmatrix} 0 & 1 \\ 1 & \alpha \end{pmatrix}$, $\alpha > 0$.

First, if $a = 0$ then it follows that $\tilde{A} = T_d$ and $d > 0$, so we are done. So assume that $a > 0$. We now claim that either $a \geq c$ and $b \geq d$ (the elements of the first row are greater than or equal to the corresponding elements of the second row) or $a \leq c$ and $b \leq d$ (the elements of the second row are greater than or equal to the corresponding elements of the first row). If this were not the case then it is impossible for the determinant of $\tilde{A}$ to be ± 1. Because $\tilde{A}$ is non-negative and has an eigenvector with slope greater than one, it must map the vector $(1 \ 1)^T$ to a vector with slope greater than one. But this means that $c + d > a + b$ so that the second pair of inequalities must hold.

Now find the largest positive integer l such that $T_l^{-1} \tilde{A}$ still has all entries non-negative. Observe that

$$T_l^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c - l \cdot a & d - l \cdot b \\ a & b \end{pmatrix} ,$$

so the inequality conditions guarantee that l can be chosen to be positive. As long as there are no zeros in the first row, the second pair of inequalities must hold since this new matrix is still an element of $Gl(2, \mathbf{Z})$ with nonnegative entries, else l was not chosen large enough. This reduction by multiplication by some T_l^{-1} may be continued, eventually reducing the matrix to either the identity or T_0.

Thus we have shown so far that one of the following must be true:

$$\begin{aligned} \tilde{A} &= T_{d_1} \cdots T_{d_p} \\ &\text{or} \\ \tilde{A} &= T_{d_1} \cdots T_{d_p} T_0 . \end{aligned}$$

However, only the first relation is possible. The second relation is eliminated by demonstrating that its right hand side does not have an eigenvector with slope greater than minus one but less than zero. We leave the details as an exercise.

We now compare the continued fraction of $\omega_{\tilde{A}}$ as computed by the product of T_{d_i} above and what we knew earlier to find that $[\overline{d_1 \ \dots \ d_p}] = [\overline{b_1 \ \dots \ b_m}]$. Since m was the (minimal) period, the $\{d_1, \dots, d_p\}$ must equal $\{b_1, \dots, b_m\}$ or some repeated sequences of $\{b_1, \dots, b_m\}$. So, $p = Nm$ and we have proved the result for $\mathrm{tr}(A) > 0$ and completed the proof of the theorem. $\square$

Remark. Effros and Shen [10] have a similar factorization result for positive matrices in $Gl(2, \mathbf{Z})$. Additionally, using a result from Hardy and Wright [13, Theorem

172] and Galois's Theorem, we can show that the factorization of $\tilde{A} = T_{d_1} \cdots T_{d_p}$ as described in the above proof is determined from the finite continued fraction of d/c from the relationship $d/c = [d_p\ \ldots\ d_1]$.

One immediate consequence of this theorem is that every quadratic surd is the slope of an eigenvector of an element of $Gl(2, \mathbf{Z})$. The results in this Theorem in addition to the following corollary to Theorem 6 allow for some interesting examples to be easily constructed.

Corollary 16. *Let $A \in \mathcal{M}^+$. Then A is conjugate to its transpose if and only if the periodic part of the continued fraction expansion of the slope of an eigenvector is a palindrome up to a circular permutation.*

Proof. The slopes of eigenvectors of a matrix A and its transpose A^* are related by the following relation:

$$\omega_{A^*} = -1/\overline{\omega}_A \ ,$$

where $\overline{\omega}$ denotes the conjugate of ω. By the remark following Theorem 10 the periodic part of the continued fraction of $-1/\overline{\omega}$ is the reverse of that of ω. Thus the equivalency above holds exactly when the periodic part is not affected by being reversed (up to a cyclic permutation). So using Theorem 6, the corollary is proven. $\square$

Example 17. The matrix $T_1T_2T_3 = \begin{pmatrix} 2 & 7 \\ 3 & 10 \end{pmatrix} \in Gl(2, \mathbf{Z})$ is not conjugate to its transpose. See also Cuntz & Krieger [7].

Example 18. The matrices $T_aT_b = \begin{pmatrix} 1 & b \\ a & ab+1 \end{pmatrix} \in Sl(2, \mathbf{Z})$, $a, b > 0$ are all conjugate to both their transposes and their inverses. (For $A \in Sl(2, \mathbf{Z})$, $\omega_{A^{-1}} = \overline{\omega}_A$, thus $A^{-1} \sim A^*$.)

Example 19. The matrix $T_1T_1T_2T_3 = \begin{pmatrix} 3 & 10 \\ 5 & 17 \end{pmatrix} \in Sl(2, \mathbf{Z})$ is similar to, but not conjugate to, its inverse.

Example 20. If $a, b > 1$, then T_1T_{ab} and T_aT_b are similar, but not conjugate.

6. The intersection quadratic form

The algebraic intersection number of two elements of the fundamental group of the torus $\pi_1(\mathbf{T}^2)$, $\begin{pmatrix} p \\ q \end{pmatrix}$ and $\begin{pmatrix} p' \\ q' \end{pmatrix}$, is given by $\begin{vmatrix} p & p' \\ q & q' \end{vmatrix}$. The topological class of a linear endomorphism f dictates how elements of $\pi_1(\mathbf{T}^2)$ intersect their image under the induced map f_* on the fundamental group. This motivates us to define the *intersection quadratic form* of a $\mathbf{Z}$-matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ by

$$\mathcal{F}_M(y, x) = \left| M \begin{pmatrix} y \\ x \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} \right| = bx^2 + (a-d)xy - cy^2 \ ,$$

where we use the unusual ordering of x and y to facilitate the comparison of our problem with the classical theory of quadratic forms. For the same reason, we also propose a skewed combination of the usual proper and improper conjugacies for quadratic forms, and say that two quadratic forms Q_1, Q_2 are *equivalent*, which we denote also by $Q_1 \sim Q_2$, if $Q_1 = \det(T) \cdot Q_2 \circ T$ for some $T \in Gl(2, \mathbf{Z})$. If $T \in Sl(2, \mathbf{Z})$, we say that two quadratic forms Q_1, Q_2 are *properly equivalent*, and we notice that this is just the classical definition due to Gauss [12]. The key to the relationship between equivalence of matrices and equivalence of forms is provided by the following result

Theorem 21. *Let M, N be similar 2×2 $\mathbf{Z}$-matrices. Then M and N are conjugate if and only if $\mathcal{F}_M \sim \mathcal{F}_N$. Furthermore the conjugacy and the equivalence are both proper or neither are.*

Proof. Suppose that $M = T^{-1}NT$, then

$$\begin{aligned}\mathcal{F}_M(y,x) &= \left| M\begin{pmatrix} y \\ x \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} \right| \\ &= \left| T^{-1}NT\begin{pmatrix} y \\ x \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} \right| \\ &= \det(T^{-1}) \left| NT\begin{pmatrix} y \\ x \end{pmatrix} T\begin{pmatrix} y \\ x \end{pmatrix} \right| \\ &= \det(T) \cdot \mathcal{F}_N \circ T(y,x)\ .\end{aligned}$$

Suppose that $\mathcal{F}_M = \det(T)\mathcal{F}_N \circ T$, then

$$\left| M\begin{pmatrix} y \\ x \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} \right| = \left| T^{-1}NT\begin{pmatrix} y \\ x \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} \right| .$$

Thus

$$\left| (M - T^{-1}NT)\begin{pmatrix} y \\ x \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} \right| = 0\ .$$

This means that the matrix $M - T^{-1}NT$ maps all lines to themselves and must be a scalar multiple of the identity. Due to similarity, the traces of M, N, and $T^{-1}NT$ are the same. So $\mathrm{tr}(M - T^{-1}NT) = 0$, and $M = T^{-1}NT$. ☐

It follows that the conjugacy and proper conjugacy problems for matrices are reduced to the classical analogous problems for quadratic forms by associating to the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ the quadratic form $Ax^2 + Bxy + Cy^2$, where

$$A = b,\ B = a - d,\ C = -c\,.$$

For instance, the reader can recognize in the formula

$$-A < B \le A < C \quad \text{or} \quad 0 \le B \le A = C\,,$$

from Lemma 8 the classical characterization of the coefficients of a reduced positive definite form. (See, *e.g.*, [8, p. 142]. Gauss [12] defined reduced positive definite quadratic forms and proved that there is exactly one in each proper equivalence class of forms. He also gave an algorithm to find it given any quadratic form in the equivalence class.)

Thus we may take as a complete invariant of conjugacy the trace of the endomorphism and the intersection form class. Note that the determinant is determined by the trace and the form.

Remarks.

1. Because traces and determinants contain together more information than the discriminant alone, matrices from different similarity classes will be associated to the same quadratic form. Every quadratic form of a fixed discriminant is associated to exactly one matrix in a fixed similarity class with the same discriminant. (Notice that the discriminant of the associated quadratic form is the same as the discriminant of the characteristic polynomial.)

2. Using this association between quadratic forms and 2×2 $\mathbf{Z}$-matrices, Sections 3-5 follow from the classical theories of quadratic forms. The principal slope corresponds to the principal root. We find that all the classical results can be proved using arguments regarding the eigenvectors and slopes which some might find easier to follow than a purely algebraic approach.

In the case of proper conjugacy, the choice of normal forms in Section 5.2.2 corresponds to Gauss's reduced indefinite quadratic forms as presented by Lejeune-Dirichlet [23].

3. Traditionally, a symmetric matrix S has been associated to a quadratic form $Q(x, y)$ so that $Q = (x\ y)S(x\ y)^T$. The equivalence relation induced on these symmetric matrices by equivalence of quadratic forms results in $S \sim CSC^T$, for $C \in Sl(2, \mathbf{Z})$.

4. The quadratic equations of (1) and (2) may now be better understood. We shall call an integer which can be represented by the intersection quadratic form associated to a matrix an *intersection number* for the matrix. Then, for the matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, b and $-c$ are intersection numbers. This gives topological meaning to the off-diagonal elements of a matrix. The necessity of the quadratic equalities in (1) and (2) now follows from the fact that conjugate matrices must have identical sets of intersection numbers, up to a sign if the conjugacy is improper.

7. Rotation numbers

For p and q coprime, let us denote by $\begin{pmatrix} p \\ q \end{pmatrix}$ both the element of $\pi_1(\mathbf{T}^2)$ and the simple closed curve on $\mathbf{T}^2$ whose lift is the vector $\begin{pmatrix} p \\ q \end{pmatrix}$. The set of translates of the curve is a (linear) foliation of $\mathbf{T}^2$ and we denote by $\Phi_{(p,q)}$ an arbitrary flow corresponding to this foliation. For a linear endomorphism f and its induced action f_* on $\pi_1(\mathbf{T}^2)$, set

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = f_* \begin{pmatrix} p \\ q \end{pmatrix} .$$

We notice that p' and q' need not be coprime if f is not an automorphism, so that $\begin{pmatrix} p' \\ q' \end{pmatrix}$ should be seen as an element of $\pi_1(\mathbf{T}^2)$ or a closed curve (not necessarily simple), covering $n = gcd(p', q')$ times a simple curve $C_{(p',q')}$. Let $F_{(p,q)}$ be the first return map of $\Phi_{(p,q)}$ on $C_{(p',q')}$. $F_{(p,q)}$ is a rotation, and we denote its rotation number by $\frac{r}{s}$, where s is the number of intersections of $\begin{pmatrix} p \\ q \end{pmatrix}$ with $C_{(p',q')}$, and r is defined mod s since it depends on the choice of a basis for $\pi_1(\mathbf{T}^2)$. One can check that the full collection $n \cdot (r + ks)$, $k \in \mathbf{Z}$, is obtained as the collection of the $\left\{ \left| f_* \begin{pmatrix} p \\ q \end{pmatrix} \begin{pmatrix} P \\ Q \end{pmatrix} \right| \right\}$ where the $\frac{P}{Q}$'s are the Farey neighbors of $\frac{p}{q}$.

Now if we choose any of these pairs (P, Q), and if $\frac{r}{s}$ is the corresponding rotation number, the matrix for f in the basis formed by $\begin{pmatrix} p \\ q \end{pmatrix}$ and $\begin{pmatrix} P \\ Q \end{pmatrix}$ has $\begin{pmatrix} n \cdot r \\ n \cdot s \end{pmatrix}$ as its first column. The matrix is thus completely determined since the trace and determinant are fixed by the similarity class. Call the pairs $(n \cdot r, n \cdot s)$ obtainable for all initial pairs (p, q) the *homotopic rotation numbers* of f (or f_*), and their collection the *rotation set* of f. It is then clear that

Theorem 22. *The rotation sets of two maps in the same similarity class are either disjoint or identical.*

One could define new normal forms on the basis of the rotation set: it seems that the best way would be to minimize s, which can be done by a theorem of Lagrange [21] by choosing $\frac{p}{q}$ among the best rational approximations of the principal slope (it is enough to look at one period of the continued fraction of this slope), and then taking $n \cdot r$ as small as possible (*i.e.*, between 0 and $n \cdot (s-1)$). While this new normal form is quite appealing conceptually, we have no associated reduction process to offer. So we conclude with the fact that the triple formed by the determinant, the trace and the rotation set (or a systematically defined element of it) forms a complete invariant of topological conjugacy.

8. Topological Interpretations

As indicated in the introduction, a matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer entries has a topological interpretation as the linear lift of a linear endomorphism f of the 2-torus onto itself; and $Gl(2, \mathbf{Z})$-conjugacy of linear lifts of two such continuous maps is equivalent to their topological conjugacy. So all the invariants ought also to have topological meaning, and they do!

As is well known, the $|\det(M)|$ is the degree of the map f; and the positivity or negativity of $\det(M)$ corresponds to whether f is orientation preserving or reversing.

The trace may be interpreted in terms of counting intersections of certain pairs of circles on the torus as follows. As mentioned in Section 6, if $u = (p\ q)^T$ and $v = (r\ s)^T$ are elements of the fundamental group $\pi_1(\mathbf{T}^2)$, then the determinant $|u\ v| = ps - qr$ is equal to the signed number of times u intersects v, the sign being determined by an orientation of elements u and v. From a simple calculation and assuming that u and v are independent, we get

$$\operatorname{tr}(M) = \frac{\left| M \begin{pmatrix} p \\ q \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix} \right| + \left| \begin{pmatrix} p \\ q \end{pmatrix} M \begin{pmatrix} r \\ s \end{pmatrix} \right|}{\begin{vmatrix} p & q \\ r & s \end{vmatrix}}$$

Thus the trace of M has the following topological meaning: it is the sum of the signed number of times u intersects the image of v under f_* with the signed number of times the image of u intersects v normalized by the signed number of times u intersects v. Another less elementary interpretation of the trace, which we shall not present, is given by the Lefschetz fixed point theorem [4].

The topological interpretation of the conjugacy invariant within a similarity class for a general **Z**-matrix M with a saddle-type phase portrait has been presented in Sections 6 and 7. In the particular case of elements of $M_F \in \mathcal{M}^+$, the periodic tail of the continued fraction expansion of the slope of an eigenvector has the following known topological interpretation (see *e.g.*, [16]; such considerations can also be traced back to [2]). Let ℓ be a ray in an eigendirection, and let u be any circle on the torus represented by a line connecting two points in $\mathbf{R}^2$ whose difference is in $\mathbf{Z}^2$. On u, fix an initial orientation. The ray ℓ intersects u a countable number of times in an ordered succession of points. Pick as origin on u the first such point. Let I be the directed sub-interval formed by the origin and the next return. Not including the origin, count the number of successive returns of ℓ to u that occur before the next one lies in I. Repeat this process: that is, form a new circle from I by identifying end-points; reverse the previous orientation; fix as origin the first return of ℓ to this new circle; mark off a new directed sub-interval with the next return; and count successive returns according to the rule above. Continue in this manner, and the sequence of numbers so obtained has a tail which coincides with that of the continued fraction expansion of the slope of the eigenvector.

Acknowledgements

PAW was a postdoctoral visitor at IBM's T. J. Watson Research Center when most of the work for this paper was done. All the authors are grateful to the referee for his careful reading of our manuscript and suggested improvements.

References

[1] R. Adler and R. Palais, *Homeomorphic conjugacy of automorphisms of the torus*, Proc. Amer. Math. Soc. **16** (1965), no. 6, 1222–1225.
[2] J. Bernoulli, Recueil pour les Astronomes (Berlin) **1** (1772), 255–284.
[3] M. Boyle and D. Handelman, *Algebraic shift equivalence and primitive matrices*, Trans. Amer. Math. Soc. **336** (1993), no. 1, 121–149.
[4] R. F. Brown, *The Lefschetz fixed point theorem*, Scott Foresman, Glenview, IL, 1971.
[5] D. A. Buell, *Binary quadratic forms: classical theory and modern computations*, Springer-Verlag, New York, 1989.
[6] H. Cohn, *A classical invitation to algebraic numbers and class fields*, Springer-Verlag, New York, 1978.
[7] J. Cuntz and W. Krieger, *Topological Markov chains with dicyclic dimension groups*, J. Reine Angew. Math. **320** (1980), 44–51.
[8] H. Davenport, *The higher arithmetic: An introduction to the theory of numbers*, 6th ed., Cambridge University Press, 1992.
[9] R. Dedekind, *Schreiben an Herrn Borchardt über die Theorie der elliptischen Modulfunktionen, Journ. f. reine u. angew. Mathem. 1877*, Mathematische Werke I (New York), Chelsea, 1969, pp. 174–201.
[10] E. G. Effros and C.-L. Shen, *Approximately finite C^*-algebras and continued fractions*, Indiana University Math. J. **29** (1980), 191–204.
[11] E. Galois, *Démonstration d'un théorème sur les fractions continues périodiques, Annales de Math. 1829*, Œvres Mathématiques d'Évariste Galois, 2nd edition (Paris), Gauthiers-Villars, 1951, pp. 1–8.
[12] K. F. Gauss, *Disquisitiones arithmeticae*, Yale University Press, New Haven, 1966, Translated by A. A. Clarke.
[13] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford, New York, 1993.
[14] G. Humbert, *Sur les fractions continues ordinaires et les formes quadratiques binaires indéfinies*, J. Math. Pure Appl. **2** (1916), 104–154.

[15] S. Katok, *Coding of closed geodesics after Gauss and Morse*, To appear in *Geometriae Dedicata.*
[16] Y. Katznelson and D. Ornstein, *The differentiability of the conjugation of certain diffeomorphisms of the circle*, Erg. Th. & Dyn. Sys. **9** (1989), 643–680.
[17] F. Klein and R. Fricke, *Vorlesungen uber die theorie der elliptischen modulfunctionen*, vol. II, B. G. Teubner, Leipzig, 1966.
[18] J. L. Lagrange, *Additions au mémoire sur la résolution des équations numériques, Nouveaux Memoires de l'Acad. Berlin, 1770*, Œvres, volume 2 (Paris), Gauthiers-Villars, 1868, pp. 603–615.
[19] ———, *Recherches d'arithmétique, Nouveaux Mémoires de l'Acad. Berlin, 1773*, Œvres, volume 3 (Paris), Gauthiers-Villars, 1869, pp. 695–758.
[20] ———, *Recherches d'arithmétique, Nouveaux Mémoires de l'Acad. Berlin, 1775*, Œvres, volume 3 (Paris), Gauthiers-Villars, 1869, pp. 759–795.
[21] ———, *Additions aux Elements d'Algèbre d'Euler: Analyse Indéterminée. St. Petersburg, 1798*, Œvres, volume 7 (Paris), Gauthiers-Villars, 1877, pp. 5–180.
[22] C. G. Latimer and C. C. MacDuffee, *A correspondence between classes of ideals and classes of matrices*, Ann. Mathematics **34** (1933), 313–316.
[23] G. Lejeune-Dirichlet, *Simplification de la théorie des formes binaires du second degré à déterminant positif, J. de Math. 1857*, Mathematische Werke II (New York), Chelsea, 1969, pp. 159–181.
[24] W. J. LeVeque, *Topics in number theory*, vol. 2, Addison-Wesley, Reading, Mass., 1956.
[25] G. B. Mathews, *Theory of numbers*, 2nd ed., Chelsea, New York, 1961.
[26] R. A. Mollin, *Quadratics*, CRC Press, Boca Raton, 1995.
[27] H. Poincaré, *Sur les courbes définies pas des équations différentielles, J. Math Pures et Appl. 4ème série 1 1885*, Œvres Complètes, t. 1 (Paris), Gauthier-Villars, Paris, 1951, pp. 90–158.
[28] J.-A. Serret, *Développements sur une classe d'équations*, J. de Math. **15** (1850), 152–168.
[29] H. J. S. Smith, *Mémoire sur les équations modulaires, Atti ar Accad. Lincei 1877*, Collected Papers, volume 2 (New York), Chelsea, 1965, pp. 224–241.
[30] ———, *Report on the theory of numbers, Part III, Report of the British Association 1861*, Collected Papers, volume 1 (New York), Chelsea, 1965, pp. 163–228.
[31] O. Taussky, *On a theorem of Latimer and MacDuffee*, Canad. J. Math. **1** (1949), 300–302.
[32] ———, *Connections between algebraic number theory and integral matrices*, H. Cohn, A Classical Invitation to Algebraic Numbers and Class Fields, Springer-Verlag, 1978.
[33] R. F. Williams, *The "DA" maps of Smale and structural stability*, Proc. Symp. in Pure Math. **AMS, vol. 14** (1970), 329–334.

I.B.M., P.O. Box 218, Yorktown Heights, NY 10598
E-mail address: adler@watson.ibm.com

I.B.M., P.O. Box 218, Yorktown Heights, NY 10598
E-mail address: tresser@watson.ibm.com

The Geometry Center, 1300 S. Second St, Minneapolis, MN 55454
E-mail address: worfolk@geom.umn.edu