On some new invariants for strong shift equivalence for shifts of finite type

Eilers, Søren; Kiming, Ian

Published in:
Journal of Number Theory

Publication date:
2012

Document Version
Early version, also known as pre-print

Citation for published version (APA):
Eilers, S., & Kiming, I. (2012). On some new invariants for strong shift equivalence for shifts of finite type. Journal of Number Theory, 132, 502-510.

ON SOME NEW INVARIANTS FOR STRONG SHIFT EQUIVALENCE FOR SHIFTS OF FINITE TYPE.

SØREN EILERS, IAN KIMING

Abstract. We introduce a new computable invariant for strong shift equivalence of shifts of finite type. The invariant is based on an invariant introduced by Trow, Boyle, and Marcus, but has the advantage of being readily computable.

We summarize briefly a large-scale numerical experiment aimed at deciding strong shift equivalence for shifts of finite type given by irreducible 2×2-matrices with entry sum less than 25, and give examples illustrating to power of the new invariant, i.e., examples where the new invariant can disprove strong shift equivalence whereas the other invariants that we use can not.

1. Introduction.

The shifts of finite type (STFs) form an important class of symbolic dynamical systems which has fundamental applications in mathematics, physics and computer science. The classification problem for irreducible SFTs up to conjugacy is generally believed to be undecidable as indicated by the examples of Kim and Roush [9] demonstrating the difference between Williams' concepts [15] of *shift equivalence* and *strong shift equivalence*. Indeed, shift equivalence is decidable [8], but it is the more elusive strong shift equivalence which encodes this significant problem, and one can no longer hope that these properties are one and the same.

When attempting to prove that two matrices **fail** to be strongly shift equivalent one has access to a large and very diverse family of invariants developed over the last decades, see [11] for a summary. Most of these invariants are efficiently computable and comparable as they take the form of algebraic numbers, finitely generated groups etc. An important invariant was developed by Trow in [14] and Boyle, Marcus, Trow in [4]. This invariant takes the form of the class of a certain ideal in a certain integral domain (see below in section 2 for details). We shall refer to it as the BMT invariant for brevity. The fact that the BMT invariant is not readily computable was the starting point and prime motivation behind the present paper.

With this invariant as basis we introduce here 2 new invariants: The first is defined and computable under a slight technical restriction and is proved to be equivalent to the BMT invariant. The second invariant is defined unconditionally, is possibly weaker than the BMT invariant, but has the advantage of being computable.

More precisely, the BMT invariant takes the form of the class (defined in the usual way) of a certain ideal of the ring $\mathbb{Z}[1/\lambda]$ where λ is a certain algebraic integer. Under a technical restriction involving the conductor of the order $O := \mathbb{Z}[\lambda]$ our first new invariant is defined as a certain element of the Picard group of O. This

new invariant is shown to be equivalent to the BMT invariant when it is defined. Our second invariant is defined unconditionally as a certain element in the class group of the algebraic number field $\mathbb{Q}(\lambda)$. This second invariant is weaker than the BMT invariant in the sense that equality of BMT invariants implies equality of our second invariants, but it is computable by standard algorithms in algebraic number theory as implemented for instance in the computer algebra package Magma, [3]. This leads to an algorithmic approach to testing this necessary condition which, as we shall see, is quite efficient in disproving strong shift equivalence where all other invariants fail.

We have combined this contribution with other tools that are already described in the literature to perform a complete analysis of the question of strong shift equivalence in a limited universe of SFTs, given by irreducible integer valued 2×2 matrices with an entry sum less than or equal to 25. Building on a project by O. Lund Jensen [10] and using standard database tools we have recorded invariants for all matrices in this universe (with the purpose of telling isomorphism classes SFTs apart) and concrete strong shift equivalences (with the purpose of identifying isomorphism classes).

The net result of these efforts can be summarized as follows. There are 17250 matrices in the universe described, and hence 148772625 potential questions of the type 'are matrices A and B strong shift equivalent?' We can answer 99.9993769 % of these questions by this approach, viz. 927 such questions are open. We will briefly summarize the methods and results of this project in section 3 below, but postpone a full description of it, along with a quantitative analysis of the efficiency of the various invariants, to [6].

2. New invariants.

Let S be an $n \times n$ matrix with non-negative, integral coefficients. We call S irreducible if for every (i, j) there is $k \geq 0$ such that the (i, j)'th entry of S^k is positive (S^0 is defined to be the identity matrix; in other words, the irreducibility condition is empty for the diagonal entries of S). Irreducibility of S corresponds to irreducibility of the associated SFT in the sense that any pair of legal words u, v can be interpolated by a third word, w, to obtain a legal word uwv.

Under these conditions, one knows, cf. for instance Theorem 4.2.3 (Perron–Frobenius Theorem) of [11], that S has a positive eigenvalue λ that occurs with multiplicity 1 in the characteristic polynomial of S and whose corresponding eigenspace is 1-dimensional, and is such that $|\mu| \leq \lambda$ for any other eigenvalue μ. Further, the eigenspace corresponding to λ is generated by a positive eigenvector, i.e., a vector with positive coordinates. This uniquely determined eigenvalue is referred to as the Perron eigenvalue of S.

If now λ is the Perron eigenvalue of S there is a corresponding eigenvector

$$\mathbf{v} = (v_1, \ldots, v_n)$$

with coordinates $v_1, \ldots, v_n$ in the ring $\mathbb{Z}[\lambda]$. As the eigenspace corresponding to λ is 1-dimensional, the vector $\mathbf{v}$ is uniquely determined up to multiplication with a non-zero element of the algebraic number field $\mathbb{Q}(\lambda)$.

We can then define the BMT invariant of S as the class $\mathcal{I}(S)$ of the ideal:

$$Rv_1 + \ldots + Rv_n$$

of the ring $R := \mathbb{Z}[1/\lambda]$. Here, 'class' has the usual meaning: Ideals C and D of R are called equivalent if there is $\xi \in \mathbb{Q}(\lambda)$ such that $\xi C = D$.

The BMT invariant is an invariant because of the following statement that follows from Theorem 12.1 in section 12.3 of [11] (see also Theorem 6.1 of [4], as well as [14]): Suppose that S and T are matrices with integral, non-negative entries that are irreducible in the above sense and have the same Perron eigenvalue. Then if the SFTs attached to S and T are strongly shift equivalent, we have $\mathcal{I}(S) = \mathcal{I}(T)$.

In Theorem 1 below we introduce 2 new invariants. The first of these is not always defined, but when it is defined for both S and T it coincides for S and T if and only if $\mathcal{I}(S) = \mathcal{I}(T)$, and is in this sense equivalent to the BMT invariant. This invariant takes values in the Picard group of the ring $\mathbb{Z}[\lambda]$.

The second invariant is always defined, takes values in the class group of the algebraic number field $\mathbb{Q}(\lambda)$, and is weaker than the BMT invariant in the sense that $\mathcal{I}(S) = \mathcal{I}(T)$ implies that the second invariants of S and T coincide.

Now let us begin to define these new invariants and prepare Theorem 1 below. We will work with the following slightly more general setup and notation:

K	:	an algebraic number field, i.e., a finite extension of $\mathbb{Q}$
O_K	:	the ring of algebraic integers in K
$\mathrm{Cl}(O_K)$	:	the class group of O_K
λ	:	an algebraic integer in K of the same degree over $\mathbb{Q}$ as K

In other words, the assumption on λ is that $K = \mathbb{Q}(\lambda)$. Then the ring

$$O := \mathbb{Z}[\lambda]$$

is of finite index in O_K, i.e., is an order of K. As a reference for the general theory of orders in algebraic number fields, see for instance Chap. 1, §12 of [12].

In particular, attached to the ring O is a certain ideal of O, the so-called conductor of O. Ideals of O prime to the conductor have unique factorizations into products of prime ideals. Attached to O is the Picard group $\mathrm{Pic}(O)$ of invertible ideals modulo principal ones; the Picard group coincides with the class group $\mathrm{Cl}(O_K)$ if $O = O_K$. The class group $\mathrm{Cl}(O_K)$ is a canonical quotient of $\mathrm{Pic}(O)$.

If C is an ideal of O_K we shall denote by $[C]$ the class of C in $\mathrm{Cl}(O_K)$; similarly, if C is an invertible ideal of O, the symbol $[C]$ denotes the class of C in $\mathrm{Pic}(O)$.

We consider the ring $\mathbb{Z}[1/\lambda]$. This ring is in fact the localization $O_{(M)}$ of O with respect to the multiplicatively closed system

$$M := \{1, \lambda, \lambda^2, \lambda^3, \ldots\} .$$

The claim follows immediately once we notice that $\lambda \in \mathbb{Z}[1/\lambda]$: For λ satisfies a polynomial equation:

$$\lambda^n + a_1\lambda^{n-1} + \ldots + a_n = 0$$

where $n := [K : \mathbb{Q}]$ and the a_i are integers. It follows that:

$$\lambda = -a_1 - \ldots - a_n \cdot \frac{1}{\lambda^{n-1}} \in \mathbb{Z}[1/\lambda] .$$

For ideals in any one of the rings we are considering above, we have the usual equivalence relation denoted by $\sim$, and defined by: $C \sim D$ if and only if there exists $\xi \in K^\times$ such that $\xi C = D$.

We will now for the remainder of this section assume that we are given two ideals $\mathfrak{A}$ and $\mathfrak{B}$ of the ring $O_{(M)}$. Since $O_{(M)}$ is a localization of O we know by general theory, see [16], Chap. IV, §10, p. 223, that every ideal of $O_{(M)}$ is extended from an ideal of O, in other words, that there are ideals A and B of O such that:

$$\mathfrak{A} = O_{(M)} \cdot A \ , \qquad \mathfrak{B} = O_{(M)} \cdot B \ .$$

We fix such ideals A and B.

If C and D are ideals of O we employ the usual notation $(C : D)$ to denote the fractional ideal:

$$(C : D) := \{\xi \in K \mid \ \xi D \subseteq C\} \ .$$

Theorem 1. *(i). Retaining the above notation, we have $\mathfrak{A} \sim \mathfrak{B}$ if and only if there are elements $x \in (A : B)$ and $y \in (B : A)$ such that:*

$$xy = \lambda^k$$

for some non-negative integer k.

(ii). Suppose that the ideals A, B, and $O \cdot \lambda$ are all prime to the conductor of O and let $\{P_i\}_{i \in I}$ be the set of prime divisors in O of the ideal $O \cdot \lambda$.

Then $\mathfrak{A} \sim \mathfrak{B}$ if and only if:

$$[A] \equiv [B] \mod \langle [P_i] \mid \ i \in I \rangle$$

in $\mathrm{Pic}(O)$.

(iii). In any case, if $\{Q_j\}_{j \in J}$ denotes the set of prime divisors in O_K of the ideal $O_K \cdot \lambda$ then a necessary condition for $\mathfrak{A} \sim \mathfrak{B}$ is that:

$$[O_K \cdot A] \equiv [O_K \cdot B] \mod \langle [Q_j] \mid \ j \in J \rangle$$

in $\mathrm{Cl}(O_K)$.

Proof. We first prove the 'only if' parts of *(i)* and *(ii)* as well as part *(iii)* simultaneously. So, suppose that $\mathfrak{A} \sim \mathfrak{B}$. Since K is the field of fractions of O there are nonzero elements $\alpha, \beta \in O$ such that $\alpha\mathfrak{A} = \beta\mathfrak{B}$, i.e., such that:

$$O_{(M)} \cdot \alpha A = O_{(M)} \cdot \beta B \ .$$

Now, O is a Noetherian ring so the ideals A and B are finitely generated O-modules. Write:

$$A = \sum_{\sigma \in S} O \cdot a_\sigma \ , \qquad B = \sum_{\tau \in T} O \cdot b_\tau$$

with certain elements $a_\sigma \in A$, $b_\tau \in B$, and finite index sets S and T.

For each $\sigma \in S$ we then have $\alpha a_\sigma \in O_{(M)} \cdot \beta B$ and so there is an element $s_\sigma \in M$ such that:

$$s_\sigma \alpha a_\sigma \in \beta B \ .$$

Similarly, there is for each $\tau \in T$ an element $t_\tau \in M$ such that:

$$t_\tau \beta b_\tau \in \alpha A \ .$$

Putting:

$$s := \prod_{\sigma \in S} s_\sigma \ , \qquad t := \prod_{\tau \in T} t_\tau \ ,$$

we conclude that:

$$s\alpha A \subseteq \beta B \ , \qquad t\beta B \subseteq \alpha A \ ,$$

and so consequently, $xB \subseteq A$ and $yA \subseteq B$ if we put:

$$x := t \cdot \frac{\beta}{\alpha} \ , \qquad y := s \cdot \frac{\alpha}{\beta} \ .$$

We have then $x \in (A : B)$, $y \in (B : A)$, and $xy = st \in M$ so that xy is a non-negative power of λ:

$$xy = \lambda^k$$

for some $k \in \mathbb{Z}_{\geq 0}$.

If now additionally the hypotheses of *(ii)* are fulfilled then the ideals A and B are invertible ideals of O. We can then write $(A : B) = AB^{-1}$, and since now

$$O \cdot x \subseteq AB^{-1}$$

we have

$$O \cdot x = AB^{-1} \cdot U$$

with a certain ideal U of O (namely, $U = A^{-1}B \cdot Ox$). Similarly,

$$O \cdot y = A^{-1}B \cdot V$$

with a certain ideal V of O. Then $UV = O \cdot xy = O \cdot \lambda^k$ which shows first that U and V are both prime to the conductor of O (since $O \cdot \lambda$ is), and then that their prime divisors are all among the prime divisors $\{P_i\}_{i \in I}$ of λ in O. Hence,

$$[U] \in \langle [P_i] \mid \ i \in I \rangle$$

in $\mathrm{Pic}(O)$; furthermore, as $O \cdot x = AB^{-1} \cdot U$, we have

$$[A] - [B] + [U] = 0$$

in $\mathrm{Pic}(O)$. We have shown the 'only if' part of *(ii)*.

For the proof of *(iii)* observe that we clearly have:

$$xO_K B \subseteq O_K A \ , \qquad yO_K A \subseteq O_K B \ .$$

Since all (nonzero) ideals of O_K are invertible, we can repeat the above arguments, substituting A and B by $O_K A$ and $O_K B$, respectively, to obtain the conclusion of *(iii)*.

Let us then prove the 'if' part of *(i)*: Suppose that we have elements $x \in (A : B)$ and $y \in (B : A)$ such that $xy = \lambda^k$ for some non-negative integer k.

As $x \in (A : B)$ we certainly have $x\mathfrak{B} = xO_{(M)} \cdot B \subseteq O_{(M)} \cdot A = \mathfrak{A}$. On the other hand, since $y \in (B : A)$ and since $xy = \lambda^k$ is a unit in $O_{(M)}$ we have:

$$\mathfrak{A} = O_{(M)} \cdot A = O_{(M)} \cdot xy \cdot A \subseteq O_{(M)} \cdot xB = x\mathfrak{B} \ .$$

Hence, $\mathfrak{A} = x\mathfrak{B}$ and $\mathfrak{A} \sim \mathfrak{B}$.

Finally, we prove the 'if' part of *(ii)*: By assumption there are then non-negative integers v_i for $i \in I$ such that:

$$(*) \qquad A \sim B \cdot \prod_{i \in I} P_i^{v_i}$$

(we can choose the v_i to be negative since $\mathrm{Pic}(O)$ is finite).

Let

$$O \cdot \lambda = \prod_{i \in I} P_i^{m_i}$$

be the prime factorization of $O \cdot \lambda$; by assumption, each m_i is nonzero. Choose then $k \in \mathbb{N}$ such that $k \cdot m_i \geq v_i$ for each i, put $u_i := k \cdot m_i - v_i$, and:

$$U := \prod_{i \in I} P_i^{u_i} \ , \qquad V := \prod_{i \in I} P_i^{v_i} \ ;$$

these are ideals of O prime to the conductor, and we have:

$$(**) \qquad UV = \prod_{i \in I} P_i^{u_i + v_i} = \prod_{i \in I} P_i^{km_i} = O \cdot \lambda^k \ .$$

Now, by $(*)$ and the definition of V we have

$$A^{-1}BV = O \cdot y$$

for some $y \in K^\times$. Since V is an ideal of O we have:

$$y \in A^{-1}BV \subseteq A^{-1}B = (B : A) \ .$$

Also, $[AB^{-1}] = [V] = [U^{-1}]$ in $\mathrm{Pic}(O)$ because of $(**)$; hence,

$$AB^{-1}U = O \cdot x$$

for some $x \in K^\times$. Since U is an ideal of O (as all u_i are ≥ 0), we see that:

$$x \in AB^{-1}U \subseteq AB^{-1} = (A : B) \ .$$

Now, $O \cdot xy = UV = O \cdot \lambda^k$ by $(**)$; changing x by a unit of O if necessary we then have $xy = \lambda^k$. By the already proved 'if' part of *(i)* we conclude that $\mathfrak{A} \sim \mathfrak{B}$. □

Remarks: In the setting of Theorem 1, the ideals A and B, as well as the fractional ideals $(A : B)$ and $(B : A)$ are all finitely generated abelian groups of rank $[K : \mathbb{Q}]$. If A and B are given explicitly via generators then generators for $(A : B)$ and $(B : A)$ can be computed.

The fractional ideals $(A : B)$ and $(B : A)$ are in particular finitely generated modules over the order O, and O-module generators can be found if the ideals A and B are given explicitly. Thus, the question of solvability of a single equation $xy = \lambda^k$ with $x \in (A : B)$ and $y \in (B : A)$ reduces to the question of solvability in the order O of a single quadratic equation

$$f(x_1, \ldots, x_s) = \lambda^k$$

where f is a quadratic form with coefficients in K that can be determined algorithmically when A and B are explicitly known.

In [7] it was remarked that the methods of that work show that there is an algorithm for deciding a question like this, i.e., the question of solvability of a quadratic equation in an explicitly given order of an algebraic number field.

Hence, condition (i) of Theorem 1 would become an algorithmically decidable criterion if one could somehow limit the k's that have to be considered to a finite number. In a sense, such a reduction to consideration of only finitely many k's is what is happening under the favorable conditions of (ii) of the theorem, the main point being the finiteness of $\mathrm{Pic}(O)$.

The question of whether condition (i) is algorithmically decidable in the general case where one or more of the ideals A, B, and $O \cdot \lambda$ are not prime to the conductor of O is a more complicated question that we will return to elsewhere.

3. The experiment.

As a part of his thesis work [10], Jensen established a database collecting systematically obtainable information concerning the classification of SFTs given by irreducible 2×2-matrices with entry sum less than 25.

In this database, strong shift equivalences were established using brute force searches computing elementary equivalences based on 2×2, 2×3, 3×2, and 3×3 matrices, supplemented by an implementation of Baker's methods ([1], [2]).

To establish lack of strong shift equivalence, the following invariants are used:

- The Jordan form of $A^\times$ (disregarding the null space of A, if necessary)
- The Bowen-Franks type groups $\mathbb{Z}^n/p(A)\mathbb{Z}^n$ where p is one of

$$x-1, x+1, 2x+1, 2x-1, x^2-x-1, x^2+x-1, x^2+2x+1,$$
$$x^2+1, x^2-1, x^2+x+1, x^2-x+1, x^2-2x+1, 2x^2-x-1,$$
$$2x^2-3x+1, 2x^2+3x+1, 2x^2+x-1, 4x^2+4x+1, 4x^2-4x+1,$$
$$4x^2-1.$$

- The invariant of Theorem 1

Perhaps surprisingly, as we shall describe in [6], it turns out that a full arsenal of polynomials p is very important for the efficiency of this approach.

Example 1. *Our database contains a multitude of examples for which the invariant introduced here is the only means of telling a pair of matrices apart, e.g.*

$$A = \begin{bmatrix} 14 & 2 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 13 & 5 \\ 3 & 1 \end{bmatrix}.$$

Indeed, the Jordan forms and all Bowen-Franks type groups are identical for these two matrices, but in the setup of Theorem 1 (iii) we get that $[O_K \cdot A] \in \langle [Q_j] \mid j \in J \rangle$ *whereas* $[O_K \cdot B] \notin \langle [Q_j] \mid j \in J \rangle$. *As a consequence of Theorem 1 (iii), the matrices are not strong shift equivalent.*

References

[1] K. A. Baker: 'Strong shift equivalence of 2×2 matrices of nonnegative integers', Ergodic Theory Dynam. Systems **3** (1983), 501–508.

[2] K. A. Baker: 'Strong shift equivalence and shear adjacency of nonnegative square integer matrices', Linear Algebra Appl. **93** (1987), 131–147.

[3] W. Bosma, J. Cannon, C. Playoust: 'The Magma algebra system. I. The user language', J. Symbolic Comput. **24** (1997), 235–265.

[4] M. Boyle, B. Marcus, P. Trow, Paul: 'Resolving maps and the dimension group for shifts of finite type', Mem. Amer. Math. Soc. **70** (1987), American Mathematical Society, 1987.

[5] R. Bowen, J. Franks: 'Homology for zero-dimensional nonwandering sets', Ann. Math. (2) **106** (1977), 73–92.

[6] S. Eilers, O. Lund Jensen, I. Kiming: 'A quantitative analysis of invaraiants for strong shift equivalence'. Technical rapport, in preparation.

[7] F. J. Grunewald, D. Segal: 'How to solve a quadratic equation in integers.' Math. Proc. Cambridge Philos. Soc. **89** (1981), 1–5.

[8] : H. Kim, F. Roush: 'Decidability of shift equivalence', Dynamical systems (College Park, MD, 1986–87), 374–424, Lecture Notes in Math. **1342**, Springer, Berlin, 1988.

[9] : H. Kim, F. Roush: 'The Williams conjecture is false for irreducible subshifts', Ann. of Math. (2) **149** (1999), 545–558.

[10] O. Lund Jensen: 'Symbolic dynamic systems and their invariants', Master Thesis, University of Copenhagen, July 2002.

[11] D. Lind, B. Marcus: 'An introduction to symbolic dynamics and coding', Cambridge University Press, Cambridge, 1995.
[12] J. Neukirch: 'Algebraic Number Theory', Grundlehren der Mathematischen Wissenschaften **322**, Springer, 1999.
[13] B. Parry, D. Sullivan: 'A topological invariant of flows on 1-dimensional spaces', Topology **14** (1975), 297–299.
[14] P. Trow: 'Resolving maps which commute with a power of the shift', Ergodic Theory Dynam. Systems **6** (1986), 281–293.
[15] R. F. Williams: 'Classification of subshifts of finite type', Ann. of Math. (2) **98** (1973), 120–153; errata, ibid. (2) **99** (1974), 380–381.
[16] O. Zariski, P. Samuel: 'Commutative algebra. Volume I.' University Series in Higher Mathematics, D. Van Nostrand Company, 1958.

Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark.
E-mail address: eilers@math.ku.dk
E-mail address: kiming@math.ku.dk