

# Decidability of membership problems for flat rational subsets of $GL(2, \mathbb{Q})$ and singular matrices

Volker Diekert 

Formale Methoden der Informatik, Universität Stuttgart, Germany  
diekert@fmi.uni-stuttgart.de

Igor Potapov 

Department of Computer Science, Ashton Building, Ashton Street, University of Liverpool, Liverpool, L69-3BX, UK  
potapov@liverpool.ac.uk

Pavel Semukhin 

Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK  
pavel.semukhin@cs.ox.ac.uk

---

## Abstract

This work relates numerical problems on matrices over the rationals to symbolic algorithms on words and finite automata. Using exact algebraic algorithms and symbolic computation, we prove various new decidability results for  $2 \times 2$  matrices over  $\mathbb{Q}$ . For that, we introduce the concept of flat rational sets: if  $M$  is a monoid and  $N$  is a submonoid, then *flat rational sets of  $M$  over  $N$*  are finite unions of the form  $L_0 g_1 L_1 \cdots g_t L_t$  where all  $L_i$ 's are rational subsets of  $N$  and  $g_i \in M$ . We give quite general sufficient conditions under which flat rational sets form an effective relative Boolean algebra. As a corollary, we obtain that the emptiness problem for Boolean combinations of flat rational subsets of  $GL(2, \mathbb{Q})$  over  $GL(2, \mathbb{Z})$  is decidable (in singly exponential time). It is possible that such a strong decidability result cannot be pushed any further inside  $GL(2, \mathbb{Q})$ .

We also show a dichotomy for nontrivial group extension of  $GL(2, \mathbb{Z})$  in  $GL(2, \mathbb{Q})$ : if  $G$  is a f.g. group such that  $GL(2, \mathbb{Z}) < G \leq GL(2, \mathbb{Q})$ , then either  $G \cong GL(2, \mathbb{Z}) \times \mathbb{Z}^k$ , for some  $k \geq 1$ , or  $G$  contains an extension of the Baumslag-Solitar group  $BS(1, q)$ , with  $q \geq 2$ , of infinite index. In the first case of the dichotomy the membership problem for  $G$  is decidable but the equality problem for rational subsets of  $G$  is undecidable. In the second case, decidability of the membership problem for rational subsets in  $G$  is open.

In the last part we prove new decidability results for flat rational sets that contain singular matrices. In particular, we show that the membership problem is decidable (in doubly exponential time) for flat rational subsets of  $\mathbb{Q}^{2 \times 2}$  over the submonoid that is generated by the matrices from  $\mathbb{Z}^{2 \times 2}$  with determinants in  $\{-1, 0, 1\}$ .

Finally, this paper improves and overarching all known decidability results for  $2 \times 2$  matrices and it also supports these results with concrete complexity bounds for the first time.

**2012 ACM Subject Classification** Theory of computation  $\rightarrow$  Formal languages and automata theory; Computing methodologies  $\rightarrow$  Symbolic and algebraic algorithms

**Keywords and phrases** Membership problem, NFA, (flat) rational set, linear group,  $GL(2, \mathbb{Q})$ ,  $GL(2, \mathbb{Z})$

## 1 Introduction

Many problems in the analysis of matrix products are inherently difficult to solve even in dimension two, and most of such problems become undecidable in general starting from dimension three or four. One of these hard questions is the *membership problem* for matrix semigroups: Given  $n \times n$  matrices  $\{M, M_1, \dots, M_m\}$ , determine whether there exist an integer  $k \geq 1$  and  $i_1, \dots, i_k \in \{1, \dots, m\}$  such that  $M = M_{i_1} \cdots M_{i_k}$ . In other words, determine whether a matrix belongs to a finitely generated (f.g. for short) semigroup. The membership problem has been intensively studied since 1947 when A. Markov showed in [32] that this problem is undecidable for matrices in  $\mathbb{Z}^{6 \times 6}$ . A natural and important generalization is the *membership problem in rational subsets* of a monoid. Rational sets are those which can be specified by regular expressions. A special case is the problem above: membership in the semigroup generated by the matrices  $M_1, \dots, M_m$ . Another difficult question is to decide the *knapsack problem*: “ $\exists x_1, \dots, x_m \in \mathbb{N}: M_1^{x_1} \cdots M_m^{x_m} = M?$ ”. Even significantly

restricted cases of these problems become undecidable for high dimensional matrices over the integers [5, 27]; and very few cases are known to be decidable, see [2, 6, 11]. The decidability of the above problems remains open even for  $2 \times 2$  matrices over integers [10, 12, 22, 26, 37].

Membership in rational subsets of  $\text{GL}(2, \mathbb{Z})$  (the  $2 \times 2$  integer matrices with determinant  $\pm 1$ ) is decidable. Indeed,  $\text{GL}(2, \mathbb{Z})$  has a free subgroup of rank 2 and of index 24 by [35]. Hence, it is a finitely generated virtually free group, and therefore the family of rational subsets forms an effective Boolean algebra [42, 45]. Two previous results extend the border of decidability for the membership problem beyond  $\text{GL}(2, \mathbb{Z})$  are [38, 39]. The first one is for semigroups of  $2 \times 2$  nonsingular integer matrices, and the second result is for  $\text{GL}(2, \mathbb{Z})$  extended by integer matrices with zero determinant.

This paper pushes the decidability border further. First of all, we consider membership problems for  $2 \times 2$  matrices over the rationals whereas [38, 39] deal only with integer matrices. Since decidability of the rational membership problem is known for  $\text{GL}(2, \mathbb{Z})$ , we focus on finitely generated (f.g. for short) subgroups  $G$  of  $\text{GL}(2, \mathbb{Q})$  which contain  $\text{GL}(2, \mathbb{Z})$ . In contrast to [38, 39] we also give concrete complexity bounds. In order to provide self contained exposition of the main results we combined a number of auxiliary results in Section 2 and Section 3. In Section 2 we characterize recognizable and rational sets in semigroups, provide techniques for transferring results for rational subsets in groups and highlight essential properties of Boolean algebras. In Section 3 we describe and prove the cubic procedure to compute the *Smith normal form* of a non-zero matrix  $g$  in  $\mathbb{Q}^{2 \times 2}$ .

In Section 4 we prove one of the main results which is Theorem 17. It states a dichotomy for a f.g. subgroup  $G$  sitting strictly between  $\text{GL}(2, \mathbb{Z})$  and  $\text{GL}(2, \mathbb{Q})$ : there are two cases which exclude each other. In the first case of the dichotomy,  $G$  is generated by  $\text{GL}(2, \mathbb{Z})$  and *finitely many* non-singular central matrices  $\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}$ . In that case  $G$  is isomorphic to  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$  for  $k \geq 1$ . It can be derived from known results in the literature about free partially commutative monoids and groups that equality test for rational sets in  $G$  is undecidable, but the membership problem in rational subsets is still decidable.

So, this is the best we can hope for groups sitting strictly between  $\text{GL}(2, \mathbb{Z})$  and  $\text{GL}(2, \mathbb{Q})$ , in general. If such a f.g. group  $G$  is not isomorphic to  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$ , then our dichotomy states that it contains a Baumslag-Solitar group  $\text{BS}(1, q)$  for  $q \geq 2$ . The Baumslag-Solitar groups  $\text{BS}(p, q)$  are defined by two generators  $a$  and  $t$  with the defining relation  $ta^p t^{-1} = a^q$ . They were introduced in [3] and widely studied since then. It is fairly easy to see (much more is known) that they have no free subgroup of finite index unless  $pq = 0$ , see [19]. As a consequence, in both cases of the dichotomy,  $\text{GL}(2, \mathbb{Z})$  has infinite index in  $G$ . Actually, we prove more, namely, if  $G$  contains a matrix of the form  $\begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix}$  with  $|r_1| \neq |r_2|$  (which is the second case in the dichotomy), then  $G$  contains some  $\text{BS}(1, q)$  for  $q \geq 2$  which has *infinite* index in  $G$ . It is wide open whether the membership to rational subsets of  $G$  can be decided in that second case. For example, let  $p \geq 2$  be a prime, and let  $G'$  be generated by  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ,  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ , and  $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ . In this case  $\begin{pmatrix} p & 0 \\ 0 & p^{-1} \end{pmatrix}$  also belongs to  $G'$ . As usual,  $\mathbb{Z}[1/p]$  denotes the ring  $\{p^n r \in \mathbb{Q} \mid n, r \in \mathbb{Z}\}$ ; and it is known by [4] that  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ,  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ , and  $\begin{pmatrix} p & 0 \\ 0 & p^{-1} \end{pmatrix}$  generate  $\text{SL}(2, \mathbb{Z}[1/p])$ . Hence,  $G'$  contains  $\text{SL}(2, \mathbb{Z}[1/p])$  as a subgroup. The structure  $\text{SL}(2, \mathbb{Z}[1/p])$  is given in [44, II.1 Cor. 2]: it is an amalgam of two copies of  $\text{SL}(2, \mathbb{Z})$  over common subgroup of finite index. It is however unknown how to decide subgroup membership in such amalgams. Moreover,  $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$  acts by conjugation on  $\text{SL}(2, \mathbb{Z}[1/p])$ , and since  $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$  generates an infinite cyclic group, we have that  $G' = \text{SL}(2, \mathbb{Z}[1/p]) \rtimes \mathbb{Z}$ . Hence, even if subgroup membership for  $\text{SL}(2, \mathbb{Z}[1/p])$  was decidable, then it could still be undecidable in  $G'$ . The situation is more friendly for the subgroup generated by the matrices  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  and  $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$  because it is the group  $\text{UT}(2, \mathbb{Z}[1/p]) \rtimes \mathbb{Z} \cong \mathbb{Z}[1/p] \rtimes \mathbb{Z} \cong \text{BS}(1, p)$ . The group  $\text{BS}(1, p)$  is metabelian and the subgroup membership is decidable in f.g. metabelian groups [40]. Actually, a stronger result is known for Baumslag-Solitar groups: decidability of membership for rational subsets in  $\text{BS}(1, q)$  for all  $q \geq 1$  was shown by Cadilhac, Chistikov, and Zetsche in [9].

In Section 7 we prove new decidability results for flat rational sets that contain singular matrices. In particular, we show that the membership problem is decidable (in doubly exponential time) for flat rational subsets of  $\mathbb{Q}^{2 \times 2}$  over the submonoid that is generated by the matrices from  $\mathbb{Z}^{2 \times 2}$  with

determinants in  $\{-1, 0, 1\}$ .

In order to keep the paper essentially self-contained and for convenience to the reader some proofs of well-known (or folklore) results have been included in the appendix, Section 9. For example the appendix contains a calculation of the generators for  $\text{SL}(2, \mathbb{Z}[1/p])$ .

## 2 Notation and preliminaries

A monoid  $M$  is a semigroup  $(M, \cdot)$  with a neutral element 1. If we use a multiplicative notation, then 1 denotes the neutral element of a monoid. In particular, the empty word in free monoids is denoted by 1 as well. A *zero* in  $(M, \cdot)$  is an element 0 such that  $x \cdot 0 = 0 \cdot x = 0$  for all  $x \in M$ . In commutative monoids without a zero-element, we might use an additive operation, and then the neutral element is denoted as 0. This is standard and there will be no risk of confusion. The set of *units* of  $M$  is the subgroup of invertible elements: it is the submonoid of  $x \in M$  such that there is some  $\bar{x} \in M$  with  $x\bar{x} = 1$ . If  $x$  is an invertible element in  $M$ , then  $x^{\mathbb{Z}}$  denotes the set of elements  $x^n$  where  $n \in \mathbb{Z}$  and  $x^{-1} = \bar{x}$ . We apply this concept when  $M$  is a monoid of  $2 \times 2$  integer matrices. In this case the group of units is  $M \cap \text{GL}(2, \mathbb{Z})$ . For a subset  $S \subseteq M$  we denote by  $\langle S \rangle$  the submonoid of  $M$  which is generated by  $S$ . Another notation for  $\langle S \rangle$  is  $S^*$ .

For groups (and more generally for monoids) we write  $N \leq M$  if  $N$  is a submonoid of  $M$  and  $N < M$  if  $N \leq M$  but  $N \neq M$ . If  $M$  is a monoid, then  $Z(M)$  denotes the *center* of  $M$ , that is, the submonoid of elements which commute with all elements in  $M$ . By  $U(M)$  we denote the group of *units* of  $M$ . It can be defined as follows:

$$U(M) = \{g \in M \mid 1 \in gM\}. \quad (1)$$

It is easy to see that  $U(M)$  is a submonoid of  $M$  (which contains 1) where every element has a right-inverse. Therefore,  $U(M)$  is a group, but not necessarily the largest subgroup (with respect to size) which appears as a subsemigroup in  $M$ .

A subsemigroup  $I$  of a monoid  $M$  is an *ideal* if  $MIM \subseteq I$ . The empty set  $\emptyset$  is an ideal. If  $M$  contains a zero, then  $\{0\}$  is the smallest nonempty ideal. If an ideal  $I$  contains an element of  $U(M)$ , then  $I = M$ . Hence,  $M \setminus U(M)$  is the greatest ideal in  $M$  which is not equal to  $M$ .

A group is finitely generated as a group if and only if it is finitely generated as a monoid.<sup>1</sup> As mentioned in the introduction write “f.g.” as an abbreviation for “finitely generated”.

Throughout we let  $\log(x) = \max\{1, \log_2(x)\}$ . Let  $f, g : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$  be two functions with values in non-negative real numbers. As usual, we let  $f \in \mathcal{O}(g)$  if there is some  $k \in \mathbb{N}$  such that  $f(n) \leq (kg(n) + k)$  for all  $n \in \mathbb{N}$ . Sometimes we measure complexities in *soft*  $\mathcal{O}$ -notation  $\tilde{\mathcal{O}}$ , too. We write  $f \in \tilde{\mathcal{O}}(g)$  if  $f \in \mathcal{O}(g \cdot \log^k(g))$  for some  $k \in \mathbb{N}$ . Thus, poly-logarithmic factors are neglected.

### 2.1 Reductions and complexity classes

We follow standard notation in complexity theory as it can be found for example in [36]. Decision problems are encoded as subsets in  $\{0, 1\}^*$ . We define complexity classes via the notion of reduction. For that we use  $\text{DTIME}(f)$  as well as  $\text{NTIME}(f)$  reductions between problems  $\mathcal{P}$  and  $\mathcal{Q}$ .

Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be some function. A  $\text{DTIME}(f)$  (*resp.*  $\text{NTIME}(f)$ ) *reduction* is realized by a deterministic (*resp.* nondeterministic) Turing machine which on every input  $u \in \{0, 1\}^*$  of length  $n$  stops after at most  $\mathcal{O}(f(n))$  steps. In particular, every computation stops. We assume that machine has a separate write-only output-tape which is initially empty. In every perhaps nondeterministic computation the contents on the output-tape is some  $v \in \{0, 1\}^*$  when the machine had stopped. By construction, we have  $|v| \in \mathcal{O}(f(n))$ .

<sup>1</sup> A nontrivial free group of rank  $k$  requires  $k$  group generators and  $k + 1$  monoid generators.

We call it a reduction from  $\mathcal{P}$  to  $\mathcal{Q}$  if for all  $u \in \mathcal{P}$  there is some (perhaps nondeterministic) computation with an output  $v$  such that  $v \in \mathcal{Q}$ . On the other hand, if  $u \notin \mathcal{P}$ , then there is no computation which can produce any output  $v$  such that  $v \in \mathcal{Q}$ .

A **P-reduction** (resp. **NP-reduction**) is a  $\text{DTIME}(f)$  (resp.  $\text{NTIME}(f)$ ) reduction where  $f$  is bounded by some polynomial. Note that the class **NP** is closed under **NP** reductions. A **DEXPTIME-reduction** (resp. **NEXPTIME-reduction**) is a  $\text{DTIME}(f)$  (resp.  $\text{NTIME}(f)$ ) reduction where  $f$  is bounded by a function of type  $2^p$  where  $p$  is some polynomial.

If there is a  $\text{DTIME}(f)$  (resp.  $\text{NTIME}(f)$ ) reduction of a problem  $\mathcal{P}$  to a singleton like  $\{1\}$ , then we say that  $\mathcal{P}$  belongs to the complexity class  $\text{DTIME}(f)$  (resp.  $\text{NTIME}(f)$ ).

## 2.2 Computation of greatest common divisors

The results of this subsection are (most probably) well-known or folklore. However, the polynomial time bounds for our main results rely on a polynomial time bound to compute binary integers  $x$  and  $y$  in Proposition 2. Elementary proofs of Lemma 1 and Proposition 2 are in the appendix.

► **Lemma 1.** *Let  $0 \neq q \in \mathbb{Z}$ . Let  $\gcd(b, d) = 1$ . Then there are  $1 \leq x, y < |q|$  such that  $\gcd(x, y) = 1$  and  $xb + yd \equiv 0 \pmod{q}$ .*

► **Proposition 2.** *Given two  $n$ -bits integer numbers  $a$  and  $b$ , we can compute in time  $\tilde{O}(n^2)$  integer numbers  $x$  and  $y$  such that  $ax + by = \gcd(a, b)$  with  $|x|, |y| \leq \max\{|a|, |b|\}$ .*

## 2.3 Recognizable and rational sets in semigroups

Throughout this subsection  $M = (M, \cdot)$  denotes a semigroup. We recall some classical facts as they can be found, for example, in the textbook of Eilenberg [17]. We do not give proofs since they are standard and they appear in Eilenberg's book, too.

► **Definition 3.** *A subset  $L \subseteq M$  belongs to the family of recognizable sets  $\text{Rec}(M)$  if there exists a homomorphism  $\varphi : M \rightarrow N$  of  $M$  to a finite semigroup  $N$  such that  $L = \varphi^{-1}(\varphi(L))$ . We also say that  $\varphi$  (resp.  $N$ ) recognizes  $L$ .*

► **Definition 4.** *The family  $\text{Rat}(M)$  has the following inductive definition using rational (aka regular) expressions.*

1.  $|L| < \infty, L \subseteq M \implies L \in \text{Rat}(M)$ .
2.  $L_1, L_2 \in \text{Rat}(M) \implies L_1 \cup L_2, L_1 \cdot L_2, \text{ and } L_1^+ \in \text{Rat}(M)$ .

Here, for  $L \subseteq M$  the set  $L^+$  denotes the subsemigroup of  $M$  which is generated by  $L$ . If  $M$  is a monoid, then the submonoid generated by  $L$  is  $L^* = L^+ \cup \{1\}$ . It is called the Kleene-star<sup>2</sup> of  $L$ .

Note that the definition of  $\text{Rat}(M)$  is intrinsic without reference to any generating set.

► **Remark 5.** Let  $G$  be any group. Then  $L \subseteq G$  is recognizable if and only if there is normal subgroup  $H$  of finite index and a finite subset  $\{g_1, \dots, g_k\} \subseteq G$  such that  $L = \bigcup \{g_i H \mid 1 \leq i \leq k\}$ . In particular, if  $G$  is infinite, then no finite subset of  $G$  is recognizable. A subgroup  $H$  belongs to  $\text{Rat}(G)$  (resp.  $\text{Rec}(G)$ ) if and only if  $H$  is f.g. (resp. the index  $G/H$  is finite), [1]. This does not hold for submonoids: the group  $\mathbb{Z} \times \mathbb{Z}$  contains the infinitely generated rational submonoid  $\{(m, n) \in \mathbb{N} \times \mathbb{N} \mid m = 0 \vee n \geq 1\} = (0, 1) + \mathbb{N} \times \mathbb{N}$ . A group  $G$  is finite if and only if  $\text{Rec}(G) = \text{Rat}(G)$  because finite subsets are rational.

<sup>2</sup> We have  $L^+ = LL^*$  and  $L^* = L^+ \cup \{1\}$ . Hence, for monoids we can replace the closure under  $L^+$  by the closure under Kleene-star without changing the class  $\text{Rat}(M)$ .

► **Proposition 6.** *Let  $h : M \rightarrow M'$  be any homomorphism of monoids. Then the following assertions hold.*

- *If  $L' \in \text{Rec}(M')$ , then  $h^{-1}(L') \in \text{Rec}(M)$ .*
- *If  $L \in \text{Rat}(M)$ , then  $h(L) \in \text{Rat}(M')$ .*
- *If  $L \in \text{Rec}(M)$  and  $K \in \text{Rat}(M)$ , then  $L \cap K \in \text{Rat}(M)$ .*
- *Kleene's Theorem, [25]: If  $\Sigma^*$  denotes a f.g. free monoid, then  $\text{Rec}(\Sigma^*) = \text{Rat}(\Sigma^*)$ .*
- *McKnight's Theorem, [33]:  $M$  is finitely generated if and only if  $\text{Rec}(M) \subseteq \text{Rat}(M)$  if and only if  $M \in \text{Rat}(M)$ .*

Thanks to Kleene's Theorem, we define for f.g. free monoids the family of *regular languages*  $\text{Reg}(\Sigma^*)$  by  $\text{Reg}(\Sigma^*) = \text{Rec}(\Sigma^*) = \text{Rat}(\Sigma^*)$ . In the present paper the notation "regular language" always refers to a subset in some finitely generated free monoid. Note that frequently  $\text{Rec}(M) \neq \text{Rat}(M)$ . This happens, for example, as soon as  $M$  is an infinite group, or if  $M$  contains a free partially commutative monoid like  $\{a, c\}^* \times \{b\}^*$ , see e. g. [15]. We shall use a weaker form of McKnight's Theorem, only: if  $M \in \text{Rat}(M)$ , then  $M$  is finitely generated. Hence, if  $M \in \text{Rat}(M)$ , then  $\text{Rec}(M) \subseteq \text{Rat}(M)$  by the third item above.

► **Definition 7.** *A nondeterministic finite automaton over subset  $S$  of  $M$  (or  $S$ -NFA for short if the semigroup  $M$  is clear from the context) is a tuple  $\mathcal{A} = (Q, \delta, I, F)$  where  $Q$  is a finite set of states,  $\delta \subseteq Q \times S \times Q$  is a finite set of transitions, and  $I, F$  are subsets of  $Q$ . The set  $I$  (resp.  $F$ ) is called the set of initial (resp. final) states.*

A transition  $(p, s, q) \in \delta$  is also written as  $p \xrightarrow{s} q$ . Let  $m \in M$ . We say that  $m$  is *accepted* by  $\mathcal{A}$  if there are  $q_0 \in I, q_n \in F$ , and  $m$  admits a factorization  $m = s_1 \cdots s_n$  such that there is a path of transitions

$$q_0 \xrightarrow{s_1} q_1 \quad \cdots \quad q_{n-1} \xrightarrow{s_n} q_n$$

with  $s_i \in S$ . For  $n = 0$  this means  $m = 1 \in M$  and  $I \cap F \neq \emptyset$ . The *accepted language*  $L(\mathcal{A})$  is the set of  $m \in M$  which are accepted by  $\mathcal{A}$ . The NFA  $\mathcal{A}$  is called *trim*, if every state is on some accepting path. Whenever convenient we assume that  $\mathcal{A}$  is trim. Note that for every NFA  $\mathcal{A}$  there is a f.g. subsemigroup  $N \leq M$  such that  $L(\mathcal{A}) \subseteq N$ . Indeed, a possible set generators for  $N$  is given by the finite set of labels of transitions in  $\delta$ .

► **Proposition 8.** *Let  $L \subseteq M$  be any subset. Then the following assertions are equivalent.*

- *We have  $L \in \text{RAT}(M)$ .*
- *There is some  $M$ -NFA  $\mathcal{A}$  such that  $L = L(\mathcal{A})$ .*
- *The set  $L$  is the image  $\varphi(K)$  of a regular set  $K \subseteq \Sigma^+$  under some homomorphism  $\varphi : \Sigma^+ \rightarrow M$ .*

## 2.4 Transfer results for rational subsets in groups

Let  $G$  be a group with a subgroup  $H$ . The aim of Section 2.4 is to show Theorem 11.<sup>3</sup> It states that  $L \subseteq H$  and  $L \in \text{Rat}(G)$  implies  $L \in \text{Rat}(H)$ . Remark 5 shows that such a statement does not hold for submonoids of a group  $G$ , but for subgroups  $H$  it holds without any hypothesis on the group  $G$  and a subgroup  $H$ . Even in this general form, our proof of Section 2.4 is conceptually simple: it is a direct transformation of a  $G$ -NFA accepting  $L$ . The general form is used in the proof of Theorem 22.

► **Lemma 9.** *Let  $H$  be a finite index subgroup of  $G$ . Then*

$$\{L \subseteq H \mid L \in \text{Rat}(G)\} = \{L \cap H \mid L \in \text{Rat}(G)\}.$$

<sup>3</sup> An independent (unpublished) proof of Theorem 11 leaning on finite transducers was given by Sénizergues [43].

**Proof.** The inclusion  $\subseteq$  is trivial. The other inclusion is clear by Proposition 6 since  $|G/H| < \infty$  implies  $H \in \text{Rec}(G)$ .  $\blacktriangleleft$

► **Remark 10.** Lemma 9 cannot be extended to the case where  $H$  has infinite index, in general. For example, the extension fails as soon  $G$  does not have the so-called *Howson property*: there are f.g. subgroups  $H, L$  such that  $L \cap H$  is not finitely generated. Free groups satisfy the Howson property, but  $G = F(a, c) \times F(b)$  does not. To see this, let  $H$  be the subgroup of (infinite index) in  $G$  which is generated by  $(a, b)$  and  $(c, 1)$ ; and let  $L$  be the subgroup of  $G$  generated by  $(a, 1)$  and  $(c, b)$ . It is enough to show that the subgroup  $G' = H \cap L$  is not rational. For that we denote by  $\pi : G \rightarrow F(a, c)$  the canonical projection onto  $F(a, c)$ . Assume by contradiction that  $G' \in \text{Rat}(F(a, b))$ . The family  $\text{Rat}(F(a, b))$  is closed under intersection by [8]. Hence,  $G' \cap a^*b^* \in \text{Reg}(\{a, b\}^*)$  since  $\{a, b\}^*$  is a free submonoid. However,  $G' \cap a^*b^* = \{a^n b^n \mid n \in \mathbb{N}\}$  is not regular. Contradiction.

► **Theorem 11.** *Let  $G$  be any group and  $H \leq G$  be a subgroup. Then*

$$\{L \subseteq H \mid L \in \text{Rat}(G)\} = \text{Rat}(H).$$

*If  $L \subseteq H$  is definable by some  $G$ -NFA  $\mathcal{A}$  with  $n$  states, then there exists an NFA  $\mathcal{A}'$  with at most  $n$  states and labels in  $H$  such that  $\mathcal{A}'$  accepts  $L$ , too.*

*Moreover, suppose first that  $G$  is a f.g. group with decidable word problem and second that the question “ $g \in H$ ?” is decidable for  $g \in G$ . Then the construction of  $\mathcal{A}'$  is effective.*

**Proof.** Without restriction, we may assume that  $\mathcal{A}$  is trim: every state  $p$  is on some accepting path. Since  $\mathcal{A}$  has only finitely many transitions, there is a finite set  $1 \in \Sigma \subseteq G$  such that each label of a transition appears in  $\Sigma$ . For every state  $p$  of  $\mathcal{A}$  we find some word  $w \in \Sigma^*$  of length less than  $n$  such that  $w_p$  labels some path from  $p$  to some final state. We let  $\Gamma$  be the union of  $\Sigma$  and the set of these words  $w_p$  viewed as elements in  $G$ . Thus, there exists a finite subset  $R \subseteq G$  with  $|R| \leq |\Gamma|$  such that, first,  $1 \in R$  and, second, the canonical mapping  $s \mapsto Hs$  from  $R$  to the set of right-cosets  $H \backslash G$  is injective. In particular, the transition are labeled by elements having the form  $as$  with  $a \in H'$  and  $s \in R$  where  $H'$  is a finite subset of  $H$ . Note that if  $G$  is a f.g. group with decidable word problem such that the question “ $g \in H$ ?” is decidable for  $g \in G$ , then we can effectively compute  $\Gamma$  and the finite set  $H'$ . Let  $G'$  the subgroup of  $G$  which is generated by the finite set  $H' \cup R$ . By construction,  $G'$  is a f.g. subgroup of  $G$  such that  $L(\mathcal{A}) \in \text{Rat}(G')$ .

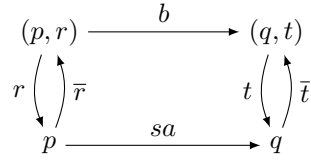
Suppose we read a word  $u$  over the alphabet  $H' \cup \Sigma$  such that reading that word from an initial state leads to the state  $p$ . Since  $\mathcal{A}$  is trim, there is some  $f \in G$  which labels a path from  $p$  to some final state. Thus,  $uf \in L(\mathcal{A}) \subseteq H$ , and therefore  $u \in Hf^{-1}$ . This means  $Hf^{-1} = Hr$  with  $r \in R$  by construction. Therefore  $r$  doesn't depend on  $u$ . It depends on  $p$  only: each state  $p \in Q$  “knows” its value  $r = r(p) \in R$  such that if  $u'$  is any word which we can read from any initial state to  $p$ , then  $u' \in Hr(p)$ . Moreover, if  $p$  is any initial or final state, then we have  $r(p) = 1$ .

Let  $r = r(p) \in R$  for  $p \in Q$ . We introduce exactly one new state  $(p, r)$  with transitions  $p \xrightarrow{\bar{r}} (p, r)$  and  $(p, r) \xrightarrow{r} p$ . This does not change the language. Recall our convention that  $\bar{r} = r^{-1}$ .

Now for each outgoing transition  $p \xrightarrow{sa} q$  with  $r = r(p)$  and  $t = r(q) \in R$  define  $b$  by the equation  $b = rsat^{-1}$ . Recall, if we read  $u$  reaching  $p$ , then  $ur^{-1} \in H$  and  $usat^{-1} \in H$ . Therefore,  $ur^{-1}rsat^{-1} \in H$  and hence  $b \in H$ . We add a transition

$$(p, r) \xrightarrow{b} (q, t).$$

This doesn't change the language as  $b = rsat^{-1}$  in  $G$  and before we added the transition there was a path  $(p, r) \xrightarrow{r} p \xrightarrow{sa} q \xrightarrow{\bar{t}} (q, t)$  as can be seen in the following picture:



Now, the larger NFA still accepts  $L$ , but the crucial point is that for  $u \in L(A)$  we can accept the same element in  $G$  by reading just labels from  $H$ . Indeed, consider any path  $p_0 \xrightarrow{s_1 a_1} p_1 \cdots \xrightarrow{s_k a_k} p_k$ , where  $k \geq 0$  and  $p_0$  is an initial. We claim that the new NFA contains a path labeled by  $b_1 \cdots b_k$  with  $b_1, \dots, b_k \in H$  from  $p_0$  to  $(p_k, r(p_k))$  such that  $b_1 \cdots b_k = s_1 a_1 \cdots s_k a_k r(p_k)^{-1}$  in  $G$ . This holds for  $k = 0$  because  $r(p_0) = 1$  and there is a transition with label 1 from  $p_0$  to  $(p_0, 1)$ . Let  $k \geq 1$ . By induction the claim holds for  $k - 1$ . Then, inspecting the figure above, where  $b = b_k$ ,  $sa = s_k a_k$ ,  $(p, r) = (p_{k-1}, r(p_{k-1}))$  and  $(q, t) = (p_k, r(p_k))$ , we see that the claim holds for  $k$  since  $r(p_{k-1})^{-1} b_k = s_k a_k r(p_k)^{-1}$ ; and so:

$$\begin{aligned}
 b_1 \cdots b_{k-1} b_k &= s_1 a_1 \cdots s_{k-1} a_{k-1} r(p_{k-1})^{-1} b_k \\
 &= s_1 a_1 \cdots s_{k-1} a_{k-1} s_k a_k r(p_k)^{-1}.
 \end{aligned}$$

We are done, since  $r(p_k) = 1$  whenever  $p_k$  is final and hence there is a transition with label 1 from  $(p_k, 1)$  to  $p_k$ .

Now we can remove all original states since they are good for nothing anymore by making  $(p, 1)$  initial (resp. final) if and only if  $p$  was initial (resp. final). Let us denote the new NFA by  $A'$ . Then  $A'$  has no more states as  $A$ .  $\blacktriangleleft$

Theorem 11 was proved first under the assumption that  $H$  has finite index in  $G$ , [20, 42, 45]<sup>4</sup>. This suffices to show the following fact.

► **Corollary 12.** *Let  $H$  be a subgroup of finite index in a f.g group  $G$ . If the membership problem for rational subsets of  $H$  is decidable, then it is decidable for rational subsets of  $G$ .*

**Proof.** Since  $H$  is of finite index, there is a normal subgroup  $N$  of finite index in  $G$  such that  $N \leq H \leq G$ . Using the canonical homomorphism from  $G$  to  $G/N$  we see that  $H$  is recognizable. Hence, “ $g \in H$ ?” is decidable. We want to decide “ $g \in R$ ?” for some  $R \in \text{Rat}(G)$ . Suppose  $u_1, \dots, u_k$  are all representatives of right cosets of  $H$  in  $G$ . Choose  $i$  such that  $g u_i^{-1} \in H$ . Then we have  $g \in R$  if and only if  $g u_i^{-1} \in R u_i^{-1} \cap H$ . Since  $H$  is recognizable, we have  $R u_i^{-1} \cap H \in \text{Rat}(G)$ . By Theorem 11, we have  $R u_i^{-1} \cap H \in \text{Rat}(H)$ ; and hence we can decide whether  $g \in R$ .  $\blacktriangleleft$

## 2.5 (Relative) Boolean algebras

► **Definition 13.** *Let  $U$  be any set and  $\mathcal{B}$  be a family of subsets of  $U$ .*

- We say that  $\mathcal{B}$  is a Boolean algebra, if  $\mathcal{B}$  is closed under finite union and complement.
- We say that  $\mathcal{B}$  is a relative Boolean algebra if  $\mathcal{B}$  is closed under finite union and relative complement:  $L, K \in \mathcal{B} \implies L \setminus K \in \mathcal{B}$ .
- We say that  $\mathcal{B}$  is an effective relative Boolean algebra if first, every  $L \in \mathcal{B}$  is given by an effective description and second, for  $L, K \in \mathcal{B}$  the union  $L \cup K$  and the relative complement  $K \setminus L$  belong to  $\mathcal{B}$ , and, in addition, an effective description for them is computable.

Every Boolean algebra is a relative Boolean algebra, and if  $\mathcal{B}$  is a relative Boolean algebra, then  $\emptyset \in \mathcal{B}$ . Moreover, a relative Boolean algebra is closed under *nonempty* finite intersection. Indeed,

<sup>4</sup> The proof in [42] states the result for f.g. virtually free groups, only.

$L \cap K = S \setminus ((S \setminus L) \cup (S \setminus K))$  where  $S = L \cup K$ . A relative Boolean algebra  $\mathcal{B}$  is a Boolean algebra if it is closed under finite intersection (that is:  $U \in \mathcal{B}$ ).

Let us give some classical examples of (relative) Boolean algebra

1. The family of regular sets  $\text{Reg}(\Sigma^*)$  is an effective Boolean algebra. More general, if  $M$  is any f.g. monoid, then the family of recognizable sets  $\text{Rec}(M)$  is an effective Boolean algebra.
2. If  $\text{Rat}(M_i)$  is an (effective) Boolean algebra for  $i = 1, 2$ , then  $\text{Rat}(M)$  is an (effective) Boolean algebra where  $M$  the free product  $M = M_1 \star M_2$ , see [41] and [29] for a generalization.
3. Let  $A$  be a f.g. abelian monoid. Then  $\text{Rat}(A)$  is an effective Boolean algebra. Rational sets in a f.g. abelian monoid are also called *semi-linear* and the result follows from [18]. In particular,  $\text{Rat}(M)$  is an effective Boolean algebra if  $M = \mathbb{Z}^k$  or  $M = \mathbb{N}^k$  for some  $k \in \mathbb{N}$ .
4. Let  $\mathbb{Q}$  be the additive group of the rational numbers. Then  $\mathbb{Q}$  is not f.g. and every f.g. subgroup is isomorphic to  $\mathbb{Z}$ . As a consequence,  $\text{Rat}(\mathbb{Q})$  is an effective relative Boolean algebra, but not a Boolean algebra.
5. A group  $G$  is called *virtually free* if it contains a free group of finite index. If  $G$  is f.g. virtually free group, then the family of rational sets  $\text{Rat}(G)$  is an effective Boolean algebra. If  $G$  is an infinitely generated free group, then  $\text{Rat}(G)$  is a relative Boolean algebra, but not a Boolean algebra. The special case of f.g. free groups is due to Michèle Benoist [8]. The extension to virtually free groups are in [20, 42, 45].

### 3 Matrices

By  $R^{n \times n}$  we denote the ring of  $n \times n$  matrices over a commutative ring  $R$ , and  $\det : R^{n \times n} \rightarrow R$  is the determinant. The units of  $R$  are denoted by  $R^*$ . We view  $R$  as a subring of  $R^{n \times n}$  by identifying  $r \in R$  with the matrix  $r \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}$ . Hence, we may write  $1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  and  $-1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ . By  $\text{GL}(n, R)$  we mean the group of invertible matrices, that is, the matrices  $g \in R^{n \times n}$  where  $\det(g) \in R^*$  is a unit. For  $n \geq 2$  the center of  $\text{GL}(n, R)$  is  $R^* = \{ \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \mid r \in R^* \}$ . By  $\text{SL}(n, R)$  we denote the *special linear group*  $\det^{-1}(1)$ . It is a normal subgroup of  $\text{GL}(n, R)$ . Of particular interest for us are  $\text{SL}(2, \mathbb{Z})$  and  $\text{GL}(2, \mathbb{Z})$ . The structure of  $\text{SL}(2, \mathbb{Z})$  is well-understood.<sup>5</sup> It is the amalgamated product  $\text{SL}(2, \mathbb{Z}) = \mathbb{Z}/4\mathbb{Z} \star_{\mathbb{Z}/2\mathbb{Z}} \mathbb{Z}/6\mathbb{Z}$ . Its quotient  $\text{PSL}(2, \mathbb{Z}) = \text{SL}(2, \mathbb{Z})/\{\pm 1\}$  is the *projective special linear group*. It is the free product  $\mathbb{Z}/2\mathbb{Z} \star \mathbb{Z}/3\mathbb{Z}$ . It is shown in [35] that  $\text{PSL}(2, \mathbb{Z})$  has a free subgroup of rank 2 and index 6. Hence,  $\text{SL}(2, \mathbb{Z})$  has a free subgroup of rank 2 of index 12 and  $\text{GL}(2, \mathbb{Z})$  has a free subgroup of rank 2 of index 24. In particular, all three groups are f.g. virtually free groups because the class of f.g. virtually free groups is closed under finite extensions.

#### 3.1 The input size of matrices and NFAs over matrices

The (bit-)complexity of an algorithm depends on the bit encoding of the input. If (in the present paper) we consider complexities, then we work mostly with NFAs where the labels of transitions are  $2 \times 2$  matrices over  $\mathbb{Q}$ . We begin with fixing the binary size of matrices. Given  $m \in \mathbb{Q}^{2 \times 2}$ , we assume that  $m$  is written as  $m = p^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  where  $p$  is the least positive integer such that  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Z}^{2 \times 2}$ . Next, we let  $\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \|_{\max} = \max\{|a|, |b|, |c|, |d|\}$  and  $\| m \|_{\max} = p \| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \|_{\max}$ . It does not yield a matrix norm, however we have:

$$\| m_1 \cdots m_n \|_{\max} \leq 2^{n-1} \prod_{i=1}^n \| m_i \|_{\max}. \quad (2)$$

<sup>5</sup> A detailed discussion about the algebraic structure of  $\text{SL}(2, \mathbb{Z})$  including the computation of normal forms can be found, for example, in [14, Sec. 8.12].



Since we are mainly interested in the bit complexity, we define  $\|m\|_{\text{bin}} = \log(\|m\|_{\text{max}})$ . In particular, for  $a, b, c, d \in \mathbb{Z}$  we have  $\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\|_{\text{bin}} = \log_2(\max\{2, |a|, |b|, |c|, |d|\})$ .

► **Lemma 14.** *Let  $m = m_1 \cdots m_n$  be a product of  $n$  matrices in  $\mathbb{Q}^{2 \times 2}$  such that  $\|m_i\|_{\text{max}} \leq 2^k$  for all  $i$ . Then we have  $\|m\|_{\text{bin}} \in \mathcal{O}(nk)$ .*

**Proof.** This is direct consequence of the inequality in (2). ◀

► **Definition 15.** *Let  $\mathcal{A} = (Q, \delta, I, F)$  be an NFA where  $\|m\|_{\text{bin}}$  is defined for all labels of transitions. The weight of a transition  $(p, m, q) \in \delta$  is defined as  $1 + \|m\|_{\text{bin}}$ . The weight  $\|\mathcal{A}\|_{\text{bin}}$  of  $\mathcal{A}$  is defined as  $1 +$  the number of states  $+$  the weight of all transitions. That is,*

$$\|\mathcal{A}\|_{\text{bin}} = 1 + |Q| + |\delta| + \sum_{(p,m,q) \in \delta} \|m\|_{\text{bin}}. \quad (3)$$

The weight  $\|\mathcal{A}\|_{\text{bin}}$  is used as the binary input size of the NFA  $\mathcal{A}$ .

### 3.2 Smith normal forms and commensurators

The intended application for our results is  $\text{GL}(2, \mathbb{Q})$ , but the results are more general. They have the potential to go beyond. Let  $n \in \mathbb{N}$ . It is a classical fact from linear algebra that each nonzero matrix  $m \in \mathbb{Q}^{n \times n}$  admits a *Smith normal form*. This is a factorization  $m = r e \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f$  such that  $r \in \mathbb{Q}$  is a positive rational number,  $e, f \in \text{SL}(n, \mathbb{Z})$ , and  $q \in \mathbb{Z}$ . The matrices  $e$  and  $f$  in the factorization are not unique, but both numbers  $r$  and  $q$  are. The existence and uniqueness of  $r$  and  $q$  are easy to see by the corresponding statement for integer matrices. More details are in Section 3.3. Clearly,  $r^2 q = \det(m)$ . So, for  $m \in \text{GL}(2, \mathbb{Q})$ , the sign of  $\det(m)$  is determined by the sign of  $q$ .

The notion of ‘‘commensurator’’ is well established in group theory. Let  $H$  be a subgroup in  $G$ , then the *commensurator* of  $H$  in  $G$  is the set of all  $g \in G$  such that  $gHg^{-1} \cap H$  has finite index in  $H$  (note that this also implies that  $H \cap g^{-1}Hg$  has finite index in  $g^{-1}Hg$ , too). If  $H$  has finite index in  $G$ , then  $G$  is always a commensurator of  $H$  because the normal subgroup  $N = \bigcap \{gHg^{-1} \mid g \in G\}$  is of finite index in  $G$  if and only if  $G/H$  is finite.

Moreover, if  $H$  has finite index in  $H'$  and if  $H' \leq G' \leq G$  such that  $G'$  is a commensurator of  $H$ , then  $G'$  is a commensurator of  $H'$ . The notion of a commensurator pops up naturally in our context. Indeed, let  $s_q = \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}$  and  $H = \text{SL}(2, \mathbb{Z})$ . Write  $g \in \text{GL}(2, \mathbb{Q})$  in its Smith normal form  $g = r e s_q f$ . Then the index of  $gHg^{-1} \cap H$  in  $H$  is the same as the index of  $s_q H s_q^{-1} \cap H$  in  $H$ ; and every matrix of the form  $\begin{pmatrix} a & b/q \\ qc & d \end{pmatrix}$  is in  $s_q H s_q^{-1}$  if  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$ . Thus, the index of  $s_q H s_q^{-1} \cap H$  in  $H$  is bounded by the size of the finite group  $\text{SL}(2, \mathbb{Z}/q\mathbb{Z})$ . For  $n = 2$  this size is in  $\mathcal{O}(q^3)$ . It follows that  $\text{GL}(2, \mathbb{Q})$  is the commensurator of  $\text{SL}(2, \mathbb{Z})$ , and hence of  $\text{GL}(2, \mathbb{Z})$ . Actually,  $\text{GL}(n, \mathbb{Q})$  is the commensurator of  $\text{SL}(n, \mathbb{Z})$  for all  $n \in \mathbb{N}$ , see for example [16, Ex. 5.19]

### 3.3 Computation of the Smith normal form

Let us recall the definition of the *Smith normal form* of a non-zero matrix  $g$  in  $\mathbb{Q}^{2 \times 2}$ . It is a factorization  $g = s/t \cdot e \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f$  where  $s$  and  $t$  are positive integers,  $e, f \in \text{SL}(2, \mathbb{Z})$ , and  $q \in \mathbb{Z}$ . Moreover,  $s/t$  and  $q$  are uniquely determined by  $g$  (but  $e$  and  $f$  are not unique). The uniqueness of  $s/t$  and  $q$  can be seen as follows. Let  $g = r_1 \cdot e_1 \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} f_1 = r_2 \cdot e_2 \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f_2$ . Multiplying  $M$  on the left by  $r_2^{-1} \cdot e_2^{-1}$  and on the right by  $f_1^{-1}$ , we see that is enough to show that that  $\frac{s}{t} \cdot e \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f$  implies  $s/t = 1$  and  $p = q$ . We may assume that  $s, t$  are positive rational numbers with  $\gcd(s, t) = 1$ . Let  $e = (e_{ij})$  and  $f = (f_{ij})$ , then

$$\begin{pmatrix} s e_{11} & s p e_{12} \\ s e_{21} & s p e_{22} \end{pmatrix} = \begin{pmatrix} t f_{11} & t f_{12} \\ t q f_{21} & t q f_{22} \end{pmatrix}.$$

Since  $\gcd(s, t) = 1$ , the positive integer  $t$  divides  $e_{11}$  and  $e_{21}$ . Hence,  $t$  divides  $\det(e) = 1$ . Thus,  $t = 1$  and by symmetry  $s = 1$ , too. Thus, we have  $e \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f$  and therefore

$\det\left(\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}\right)$ . Clearly, this implies  $p = q$ . The following lemma is a special case of result by Kannan and Bachem, [24]. We include a proof because the result for  $2 \times 2$  matrices is rather easy to show. This allows us to keep the paper essentially self-contained. Moreover, for  $2 \times 2$  matrices we obtain a soft cubic time bound which might be better than the polynomial time bound for the general case.<sup>6</sup>

► **Lemma 16.** *On input  $m \in \mathbb{Q}^{2 \times 2}$  with  $n = \|m\|_{\text{bin}}$  we can compute  $r \in \mathbb{Q}$ ,  $e, f \in \text{SL}(2, \mathbb{Z})$ , and  $q \in \mathbb{Z}$  in polynomial time  $\tilde{O}(n^3)$  such that  $m = r \cdot e \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f$ .*

**Proof.** Our proof follows the same lines as in [24]. The assertion is trivial, if  $m$  is the zero matrix. So let  $m \neq 0$ . On input  $m$  we calculate the smallest positive integer  $p$  such that  $g = p \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  where  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Z}^{2 \times 2}$ . Hence, w.l.o.g. we may assume that  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  and  $a = \|A\|_{\text{max}} > 0$ . In the following  $g, g'$  denote divisors of  $a$ , and  $D = \det(A)$ .

First step: If  $a = 1$ , then we are done because

$$\begin{pmatrix} 1 & 0 \\ -c & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & D \end{pmatrix}.$$

Hence, from now on we assume  $a \geq 2$ .

Let  $g = \gcd(a, b) = pa + qb$  with  $0 \leq q < a$ . This is possible since  $(p+b)a + (q-a)b = pa + qb$ . Then  $\begin{pmatrix} p & -b/g \\ q & a/g \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$ . Hence:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} p & -b/g \\ q & a/g \end{pmatrix} = \begin{pmatrix} g & 0 \\ pc+qd & D/g \end{pmatrix}.$$

If  $\gcd(a, b) = a$ , then we choose  $p = 1$  and  $q = 0$ . Otherwise  $a/g \geq 2$  and  $|b| < a = \|A\|_{\text{max}}$ . Hence:

$$|p| = \left| \frac{g - qb}{a} \right| \leq \frac{g}{a} + \frac{(a-1)|b|}{a} \leq 1/2 + a - 1 < a.$$

Hence, after the first step and by left-right symmetry due to transposition of matrices, we may assume without restriction that we actually start with a matrix

$$A' \in \left\{ \begin{pmatrix} g & 0 \\ D' & D/g \end{pmatrix}, \begin{pmatrix} g & D' \\ 0 & D/g \end{pmatrix} \right\}$$

where  $g|a$  and  $0 \leq |D'| < 2\|A\|_{\text{max}}^2$ . If  $D' = 0$ , then we stop because the matrix is diagonal which is the aim for this phase.

Otherwise, let  $g' = \gcd(g, D') = pg + qD'$  with  $0 \leq q < g$ . Hence

$$\begin{pmatrix} p & q \\ -D'/g' & g/g' \end{pmatrix} \cdot \begin{pmatrix} g & 0 \\ D' & D/g \end{pmatrix} = \begin{pmatrix} g' & qD/g \\ 0 & D/g' \end{pmatrix}.$$

Clearly:  $g'|g|a$ . Let  $D'' = qD/g$ . Then

$$0 \leq |D''| < |D| \leq 2\|A\|_{\text{max}}^2.$$

If  $g|D'$ , then we choose  $p = 1$  and  $q = 0$ . Then  $g' = g$  and we obtain  $\begin{pmatrix} g & 0 \\ 0 & D/g \end{pmatrix}$  which is diagonal and stops the process. (This happens in particular, if  $g = 1$ .)

For  $g|D'$  we have  $q = 0$  and  $|p| = 1$ . Otherwise, since  $g' \leq |D'|$ , we have

$$|p| = \left| \frac{g' - qD'}{g} \right| \leq \frac{g'}{g} + \frac{(g-1)|D'|}{g} = |D'| + \frac{g'}{g} - \frac{|D'|}{g} \leq |D'| < 2\|A\|_{\text{max}}^2.$$

Since either each time  $g/g' \geq 2$  or  $g|D'$ , we finish after at most  $\log \|A\|_{\text{max}}$  steps.

<sup>6</sup> We do not know the current state of the art with respect to the time complexities for the result in [24].

We continue with a matrix  $A'' = \begin{pmatrix} g & 0 \\ 0 & D/g \end{pmatrix}$  for some  $g|a$ . If  $g|D/g$  we are done. Thus, w.l.o.g.  $D \neq 0$  and letting  $d = D/g$  we write

$$\begin{pmatrix} g & 0 \\ 0 & d \end{pmatrix} = \begin{pmatrix} \gcd(g,d) & 0 \\ 0 & \gcd(g,d) \end{pmatrix} \cdot \begin{pmatrix} g/\gcd(g,d) & 0 \\ 0 & d/\gcd(g,d) \end{pmatrix}.$$

Let  $g' = g/\gcd(g,d)$  and  $d' = d/\gcd(g,d)$ . Note that  $g' \geq 2$  because  $g \neq \gcd(g,d)$ . We add the right column of  $\begin{pmatrix} g' & 0 \\ 0 & d' \end{pmatrix}$  to the left one by multiplying with the matrix  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . We obtain the matrix  $\begin{pmatrix} g' & 0 \\ d' & d' \end{pmatrix}$ . We let  $pg' + qd' = 1$  with  $0 \leq q < g'$  and  $p = (1 - qd')/g'$ . Hence,  $|p| \leq |d'| + 1/g' - |d'|/g' \leq |d'| \leq |D| \leq 2\|A\|_{\max}^2$ . Then,

$$\begin{pmatrix} p & q \\ -d' & g' \end{pmatrix} \cdot \begin{pmatrix} g' & 0 \\ d' & d' \end{pmatrix} = \begin{pmatrix} 1 & qd' \\ 0 & g'd' \end{pmatrix}.$$

Subtracting  $qd'$  times the left column from the right one by multiplying with the matrix  $\begin{pmatrix} 1 & -qd' \\ 0 & 1 \end{pmatrix}$ , we obtain the desired result.  $\blacktriangleleft$

We summarize. For every non-zero matrix  $m = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Z}^{2 \times 2}$  we can calculate in soft cubic time  $r = \gcd\{a, b, c, d\}$  (hence a divisor of  $\|m\|_{\max}$ ) and matrices  $e, f \in \text{SL}(2, \mathbb{Z})$  such that  $m = r e \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} f$ . Moreover,  $r$  and  $q$  are uniquely defined by  $m$ .

#### 4 Dichotomy in $\text{GL}(2, \mathbb{Q})$

One of the main results is Theorem 17. It states a dichotomy for a f.g. subgroup  $G$  sitting strictly between  $\text{GL}(2, \mathbb{Z})$  and  $\text{GL}(2, \mathbb{Q})$ : there are two cases which exclude each other. Moreover,  $\text{Rat}(G)$  is not closed under finite intersection in both cases. In particular,  $\text{Rat}(G)$  is not a relative Boolean algebra. In the dichotomy the Baumslag-Solitar group  $\text{BS}(1, q) = \langle a, t \mid ta^p t^{-1} = a^q \rangle$  for  $q \geq 2$  pops up. The  $\text{BS}(1, q)$  has a faithful representation in  $\text{SL}(2, \mathbb{Z})$ . It belongs to the widely studied family of *Baumslag-Solitar groups*  $\text{BS}(p, q) = \langle a, t \mid ta^p t^{-1} = a^q \rangle$ . In algebraic terms  $\text{BS}(1, q)$  is an HNN-extension of  $\mathbb{Z}$  over the isomorphism between  $\mathbb{Z}$  and its subgroup  $q\mathbb{Z}$  which sends 1 to  $q$ .

**► Theorem 17.** *Let  $G$  be a f.g. group such that  $\text{GL}(2, \mathbb{Z}) < G \leq \text{GL}(2, \mathbb{Q})$ . Then  $\text{Rat}(G)$  is not closed under finite intersection. Moreover, there are two mutually exclusive cases.*

1.  $G$  is isomorphic to  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$  for some  $k \geq 1$ .
2.  $G$  contains a subgroup which is an extension of infinite index of  $\text{BS}(1, q)$  for some  $q \geq 2$ .

**Proof.** We first show the dichotomy. Then we give a direct proof showing that  $\text{Rat}(G)$  is never closed under finite intersection if  $\text{BS}(1, q) \leq G$  and  $q \geq 2$ . Finally, we show the analogous statement for groups containing  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}$ .

Let  $H = \text{GL}(2, \mathbb{Z})$ . There are two cases. In the first case some finite generating set for  $G$  contains only elements from  $H$  and from the center  $Z(G)$ . Since  $\text{GL}(2, \mathbb{Z}) \leq G$  we see that  $Z(G) \leq \{ \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \mid r \in \mathbb{Q} \}$ . Moreover, since  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in H$ , we may assume in the first case that  $G$  is generated by  $H$  and f.g. subgroup  $Z \leq \{ \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \mid r \in \mathbb{Q} \wedge r > 0 \}$ . The homomorphism  $g \mapsto |\det(g)|$  embeds  $Z$  into the torsion free group  $\{r \in \mathbb{Q}^* \mid r > 0\}$ . Hence,  $Z$  is isomorphic to  $\mathbb{Z}^k$  for some  $k \geq 1$ . Since  $Z \cap H = \{1\}$ , the canonical surjective homomorphism from  $Z \times H$  onto  $G$  is an isomorphism.

In the second case we start with any generating set and we write the generators in Smith normal form  $e \begin{pmatrix} r & 0 \\ 0 & rq \end{pmatrix} f$ . Since  $e, f \in \text{GL}(2, \mathbb{Z})$  and  $\text{GL}(2, \mathbb{Z}) < G$ , without restriction, the generators are either from  $\text{GL}(2, \mathbb{Z})$  or they have the form  $\begin{pmatrix} r & 0 \\ 0 & rq \end{pmatrix}$  with  $r > 0$  and  $0 \neq q \in \mathbb{N}$ . So, if we are not in the first case, there is at least one generator  $s = \begin{pmatrix} r & 0 \\ 0 & rq \end{pmatrix}$  where  $r > 0$  and  $2 \leq q \in \mathbb{N}$ .

Let  $\text{BS}$  be the subgroup of  $G$  which is generated by  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  and  $s$ ; and let  $\text{BS}(1, q)$  be the Baumslag-Solitar group with generators  $b$  and  $t$  such that  $tbt^{-1} = b^q$ . We have  $s \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} s^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^q$ . Hence,

there is a surjective homomorphism  $\varphi : \text{BS}(1, q) \rightarrow \text{BS}$  such that  $\varphi(t) = s$  and  $\varphi(b) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ . Let us show that  $\varphi$  is an isomorphism. Every element  $g \in \text{BS}(1, q)$  can be written in the form  $t^k b^x t^n$  where  $k, x, n$  are integers. Suppose  $g = t^k b^x t^n$  and  $\varphi(g) = 1$ . Then  $\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = \varphi(b^x) = \varphi(t^{-k-n}) = \begin{pmatrix} r & 0 \\ 0 & rq \end{pmatrix}^{-k-n}$  is a diagonal matrix and  $x = 0$ . Hence,  $g = t^m$  and  $\varphi(g) = s^m = 1$ . This implies  $m = 0$ , and  $\varphi$  is an isomorphism and  $\text{BS}$  is the group  $\text{BS}(1, q)$ .

Next, consider any  $g \in \text{BS} \cap \text{SL}(2, \mathbb{Z})$ . As above  $g = s^k \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^x s^m$  with  $x, k, m \in \mathbb{Z}$ . Since by assumption  $\det(g) = 1$  we obtain  $m = -k$  and hence  $g = \begin{pmatrix} 1 & 0 \\ q^k x & 1 \end{pmatrix} \in \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\mathbb{Z}}$ . Therefore  $\text{SL}(2, \mathbb{Z}) \cap \text{BS}$  is the infinite cyclic group generated by  $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ . It has infinite index in  $\text{SL}(2, \mathbb{Z})$ . It follows that  $G$  contains an extension of  $\text{BS}(1, q)$  of infinite index.

But this is not enough, we need to show that  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$  cannot contain  $\text{BS}(1, q)$ , otherwise there is no dichotomy. Actually, we do more: there is no abelian group  $A$  such that  $\text{BS}(1, q)$  is a subgroup of  $\text{GL}(2, \mathbb{Z}) \times A$ .

Assume by contradiction that it is. Then there are generators  $b = (a, x), t = (s, y) \in \text{GL}(2, \mathbb{Z}) \times A$  such that  $btb^{-1} = b^q$ . This implies  $(q-1)x = 0$ . Since  $q \geq 2$ , the element  $x$  generates a finite subgroup in  $A$ . Since  $b$  generates an infinite cyclic group, we conclude that  $a^m \neq 1$  for all  $m \neq 0$ . Consider the canonical projection  $\varphi$  of  $\text{GL}(2, \mathbb{Z}) \times A$  onto  $\text{GL}(2, \mathbb{Z})$  such that  $\varphi(b) = a$  and  $\varphi(t) = s$ . We claim that the restriction of  $\varphi$  to  $\langle b, t \rangle$  is injective.

Let  $\varphi(g) = 1$  for  $g \in \langle b, t \rangle$ . As above we write  $g = t^k b^z t^n$  with  $z, k, n \in \mathbb{Z}$ . Then we have  $s^k a^z s^n = 1 \in \text{GL}(2, \mathbb{Z})$ ; and therefore  $a^z = s^{-k-n}$ . Hence  $a^z$  commutes with  $s$ . Hence  $a^z = sa^z s^{-1} = a^{qz}$ . We conclude  $a^{(q-1)z} = 1$ . Since  $a^m \neq 1$  for all  $m \neq 0$  and  $q \geq 2$  we have  $z = 0$ . Hence  $g = t^m$  for some  $m \in \mathbb{Z}$ . Since  $\varphi(g) = 1$ , we know  $s^m = 1$ . Therefore,  $t^m = (s^m, my)$  acts trivially on  $b$ . But in  $\text{BS}(1, q)$  this happens for  $m = 0$ , only. This tells us that  $\varphi$  is injective on  $\langle b, t \rangle$ , and the claim follows.

The above claim implies that  $\text{BS}(1, q)$  appears as a subgroup in  $\text{GL}(2, \mathbb{Z})$  but  $\text{GL}(2, \mathbb{Z})$  is virtually free. Hence it is hyperbolic. This contradicts a result of Gersten. His paper [19] shows that a Baumslag-Solitar group  $\text{BS}(p, q)$  with  $pq \neq 0$  cannot appear in any hyperbolic group. However, instead of using [19], let us derive a contradiction by showing a stronger result. We claim that  $\text{Rat}(\text{BS}(1, q))$  (for  $q \geq 2$ ) is not closed under intersection.<sup>7</sup> This leads to a contradiction because  $\text{Rat}(\text{BS}(1, q)) \subseteq \text{Rat}(\text{GL}(2, \mathbb{Z}))$  and  $\text{Rat}(\text{GL}(2, \mathbb{Z}))$  is a Boolean algebra.

Recall that  $\text{BS}(1, q)$  has two generators  $a$  and  $t$  with one defining relation  $ta\bar{t} = a^q$  where,  $\bar{x} = x^{-1}$  for  $x \in \text{BS}(1, q)$ . For the ease of notation, assume  $q = 2$ . We view  $\{a, \bar{a}, t, \bar{t}\}^*$  as a free monoid of rank 4 and  $\psi : \{a, \bar{a}, t, \bar{t}\}^* \rightarrow \text{BS}(1, q)$  denotes the canonical projection. Now, consider  $L = t^* a \bar{t}^* \cap a^*$ . Then  $L = \{a^{2^n} \mid n \in \mathbb{N}\}$  is the intersection of two rational sets. Assume by contradiction that  $L$  was rational. Then  $L = \psi(K)$  for some regular language  $K \subseteq \{a, \bar{a}, t, \bar{t}\}^*$ . By the pumping-lemma for regular languages (also known as  $uvw$ -Theorem), there is a constant  $p$  such that whenever  $z \in K$  with  $|z| > p$ , we can factorize  $z = uvw$  such that  $|uv| \leq p$ ,  $|v| \geq 1$ , and  $uw \in K$ . Consider  $a^{2^n} \in L$  such that its  $K$ -geodesic length  $m$  (= length of a shortest representing word in  $K$ ) is larger than  $p$ . Choose  $z \in K$  such that  $\psi(z) = a^{2^n}$  and  $|z| = m$ . Factorize  $z = uvw$  such that  $|uv| \leq p$ ,  $|v| \geq 1$ , and  $uw \in K$ . Let  $g = \psi(uw)$  and  $h = a^{2^n} = \psi(uvw)$ . Note that  $g \neq h$  because  $z$  is the smallest representative in  $K$ . Since  $uw \in K$ , we have  $g = a^{2^k}$  for some  $k \neq n$ . Assume  $k < n$  as the other case is easier. Consider  $hg^{-1} = \psi(uv\bar{u}) = a^{2^n - 2^k} = a^{2^k(2^{n-k} - 1)}$ . Note that  $|uv\bar{u}| \leq 2|p|$ , and hence there are finitely many possible values for  $\psi(uv\bar{u})$ . Therefore,  $k$  and  $n_k$  must be bounded by some constant  $C$ , which implies that  $n \leq 2C$ . This gives a contradiction since  $n$  can be arbitrarily large. Hence, the claim: if  $\text{BS}(1, q) \leq G$ , then  $\text{Rat}(G)$  is not closed under intersection, and therefore  $G \neq \text{GL}(2, \mathbb{Z})$ .

Finally,  $\text{GL}(2, \mathbb{Z})$  contains a free subgroup of rank 2. Hence, for  $k \geq 1$  the group  $G = \text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$  contains the free partially commutative monoid  $M = \{a, b\}^* \times \{c\}^*$ . In Remark 10 we have

<sup>7</sup> This fact is also stated in [9], but with a different proof.

seen that  $\text{Rat}(G)$  is not closed under intersection. ◀

► **Proposition 18.** *Let  $G$  be isomorphic to  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$  with  $k \geq 1$ . Then, the question “ $L = R?$ ” on input  $L, R \in \text{Rat}(G)$  is undecidable. However, the question “ $g \in R?$ ” on input  $g \in G$  and  $R \in \text{Rat}(G)$  is decidable.*

**Proof.** As just stated in the last lines of the proof above,  $G$  contains the free partially commutative monoid  $M = \{a, b\}^* \times \{c\}^*$ . It was by Aalbersberg and Hooeboom in [23] that the question “ $L = R?$ ” on input  $L, R \in \text{Rat}(G)$  is undecidable for  $M$ .

For the decidability we use the fact that  $\text{GL}(2, \mathbb{Z})$  has a free subgroup  $F$  of rank two and index 24. In [30] Lohrey and Steinberg showed that the question “ $g \in R?$ ” is decidable for  $F \times \mathbb{Z}^k$ . Since  $F \times \mathbb{Z}^k$  is of finite index (actually 24) in  $G$ , the membership problem for rational subsets in  $G$  is decidable by Corollary 12. ◀

► **Remark 19.** Let  $G$  be a group extension of  $\text{GL}(2, \mathbb{Z})$  inside  $\text{GL}(2, \mathbb{Q})$  which is not isomorphic to  $\text{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$  for  $k \geq 0$ . Then, by Theorem 17, the group  $G$  contains an infinite extension of  $\text{BS}(1, q)$  for  $q \geq 2$ . By [9] the membership in rational sets of  $\text{BS}(1, q)$  is decidable. However, to date it is not clear how to extend this result to infinite extensions of  $\text{BS}(1, q)$ .

## 5 Flat rational sets

In Definition 20 we introduce the general notion of a *flat rational set*. It allows to extend decidability results for  $\text{Rat}(N)$  to a larger family  $\text{Frat}(M, N)$  whenever  $N \leq M$ . The main result in this section is Theorem 23. It implies that the membership problem and (even stronger) the emptiness problem for Boolean combinations of flat rational sets are decidable in  $\text{Frat}(\text{GL}(2, \mathbb{Q}), \text{GL}(2, \mathbb{Z}))$ .

The best situation is when  $\text{Rat}(M)$  is an effective Boolean algebra because in this case all decision problems we are studying here are decidable. However, our focus is on matrices over the rational or integer numbers, in which case such a strong assertion is either not known or wrong. The most prominent example is the direct product  $F_2 \times F_2$  of two free groups of rank 2 in which, due to the construction of Mihailova [34], there exists a finitely generated subgroup with undecidable subgroup membership problem.

The notation is as follows. Let  $M = (M, \cdot)$  be a semigroup containing a submonoid  $N$  with a subset  $T$  such that  $N$  is the subsemigroup generated by  $T$ . That is, we can write  $N = \langle T \rangle$ .

► **Definition 20.** *We say that  $L \subseteq M$  is a flat rational subset of  $M$  over  $T$  if  $L$  is a finite union of languages of the form  $L_0 g_1 L_1 \cdots g_t L_t$  where all  $L_i \in \text{Rat}(\langle T \rangle)$  and  $g_i \in M$ . The family of these sets is denoted by  $\text{Frat}(M, T)$  or  $\text{Frat}(M, N)$  since  $N = \langle T \rangle$ .*

In order to specify a set in  $\text{Frat}(M, T)$  we shall use an  $M$ -NFA with a syntactic restriction according to Definition 21. In particular, as soon as membership to  $T$  is decidable, we can check whether an  $M$ -NFA is flat over  $T$ , and, if yes, this entails that the accepted language belongs to  $\text{Frat}(M, T)$ .

► **Definition 21.** *An  $M$ -NFA  $\mathcal{A} = (Q, \delta, I, F)$  is called flat over a subset  $T \subseteq M$  if no transition labeled by an element  $m \in M \setminus T$  lies on a directed cycle.*

Let  $M$  be any semigroup containing a subgroup  $H$ . Then the neutral element in  $H$  is an idempotent  $e = e^2$ . The subsemigroup  $eMe$  of  $M$  is a monoid and  $H$  is a subgroup of its units  $U(eMe)$ . Recall that  $M \setminus H$  is an ideal of  $M$  if and only if  $H = U(eMe)$ . This fact can be easily derived from Equation (1). In such a situation Theorem 22 yields an inductive definition of  $\text{Frat}(M, H)$  as a subset of  $\text{Rat}(M)$ : Theorem 22 says that the class of flat rational sets of  $M$  over  $H$  can be defined as the family of rational sets when the Kleene-star is restricted to subsets which belong to the submonoid  $H$ .

► **Theorem 22.** *Let  $M$  be a semigroup containing a subgroup  $H$  with neutral element  $e$ . Then the family  $\text{Frat}(M, H)$  is the smallest family  $\mathcal{R}$  of subsets of  $M$  satisfying the following conditions.*

- $\mathcal{R}$  contains all finite subsets of  $M$ ,
- $\mathcal{R}$  is closed under finite union and concatenation,
- $\mathcal{R}$  is closed under taking the Kleene-star over subsets of  $H$  which belong to  $\mathcal{R}$ .

**Proof.** As usual, let  $U(eMe)$  be the group of units of the monoid  $eMe$ . For the proof let  $G = U(eMe)$ . Clearly,  $\text{Rat}(H) \subseteq \mathcal{R}$  and hence, all flat rational sets over  $H$  are contained in  $\mathcal{R}$ . To prove inclusion in the other direction, we need to show that the family of flat rational subsets of  $M$  over  $H$  (i) contains all finite subsets of  $M$ , (ii) is closed under finite union and concatenation, and (iii) is closed under taking the Kleene-star over subsets of  $H$ . The first two conditions are obvious. We show (iii) in two steps. Let  $L$  be a flat rational set over  $H$  such that  $L \subseteq H$ . In the first step we show  $L \in \text{Rat}(G)$ . Recall that  $L$  is a finite union of languages  $L_0 g_1 L_1 \cdots g_t L_t$ , where  $\emptyset \neq L_i \in \text{Rat}(H)$  and  $g_i \in M$ . If  $g_i \in M \setminus G$  for some  $i$ , then we have  $L_0 g_1 L_1 \cdots g_t L_t \setminus G \neq \emptyset$  because  $M \setminus G$  is an ideal. Hence,  $L \not\subseteq G$ . Thus, if  $L \subseteq H$ , then all  $g_i \in G$  and  $L \in \text{Rat}(G)$ . In the second step we apply Theorem 11. It implies that  $L$  is a rational subset of  $H$ , and hence  $L^* \in \text{Rat}(H)$ . In particular,  $L^*$  is flat rational over  $H$ . ◀

In our applications of Theorem 22 we have  $M = \mathbb{Q}^{2 \times 2}$  to be the multiplicative monoid of  $2 \times 2$  matrices with rational entries and  $H = \text{GL}(2, \mathbb{Z})$ . In this setting it is clear that  $\mathbb{Q}^{2 \times 2} \setminus \text{GL}(2, \mathbb{Q})$  is an ideal. As we will see the following theorem applies to  $H = \text{GL}(2, \mathbb{Z})$  and  $H \leq G \leq \text{GL}(2, \mathbb{Q})$  where  $G$  is finitely generated.

► **Theorem 23.** *Let  $H$  be a subgroup of a group  $G$  with decidable word problem<sup>8</sup> such that the following conditions hold:*

- $\text{Rat}(H)$  is an effective relative Boolean algebra.<sup>9</sup>
- $G$  is the commensurator of  $H$ , and on input  $g \in G$ , we can compute the index of  $H_g$  in  $H$ .
- On input  $g \in G$ , the membership to  $H$  (that is, “ $g \in H$ ?”) is decidable.

*Then  $\text{Frat}(G, H)$  forms an effective relative Boolean algebra. In particular, given a finite Boolean combination  $B$  of flat rational sets of  $G$  over  $H$ , we can decide the emptiness of  $B$ .*

Before giving the proof of Theorem 23 let us first state a consequence.

► **Corollary 24.** *Let  $B \subseteq \text{GL}(2, \mathbb{Q})$  be a finite Boolean combination of flat rational sets of  $\text{GL}(2, \mathbb{Q})$  over  $\text{GL}(2, \mathbb{Z})$ , then we can decide the emptiness of  $B$ .*

**Proof.** As explained above, it is a well-known classical fact that  $\text{GL}(2, \mathbb{Z})$  is a finitely generated virtually free group, namely, it contains a free subgroup of rank 2 and index 24. Hence  $\text{Rat}(\text{GL}(2, \mathbb{Z}))$  is an effective Boolean algebra by [45]. Let  $G$  be a f.g. subgroup of  $\text{GL}(2, \mathbb{Q})$  that contains  $B$ . Clearly,  $G$  has a decidable word problem. Section 3.2 states that  $\text{GL}(2, \mathbb{Q})$  is the commensurator subgroup of  $\text{GL}(2, \mathbb{Z})$  in  $\text{GL}(2, \mathbb{Q})$ . Hence,  $G$  is the commensurator of  $\text{GL}(2, \mathbb{Z})$ , too. Thus all hypotheses of Theorem 23 are satisfied. ◀

A direct consequence of Corollary 24 is that we can decide the membership in flat rational subsets of  $\text{GL}(2, \mathbb{Q})$  over  $\text{GL}(2, \mathbb{Z})$ . However, by Section 4, we are far away to decide the membership for all rational subsets of  $\text{GL}(2, \mathbb{Q})$ . It is tempting to believe that  $\text{Rat}(\text{GL}(2, \mathbb{Q}))$  has an undecidable membership problem.

For the proof of Theorem 23 we need the following observation.

<sup>8</sup> In case  $G$  is not f.g., we assume that  $G$  comes with an effective presentation. For example:  $G \leq \text{GL}(n, \mathbb{Q})$ .

<sup>9</sup> Recall that this does not imply  $H \in \text{Rat}(H)$ : possibly  $H$  is not f.g.

► **Lemma 25.** *Let  $L \in \text{Rat}(H)$  and  $g \in G$ . Recall that*

$$H_g = gHg^{-1} \cap H = \{h \in H \mid g^{-1}hg \in H\}.$$

*Then under the assumptions of Theorem 23 we can compute an NFA accepting  $g^{-1}(L \cap H_g)g \in \text{Rat}(H)$  where all labels of transitions are in  $H$ .*

**Proof.** Since  $H_g = gHg^{-1} \cap H$  is of finite index in  $H$ , we can compute an NFA  $\mathcal{A}'$  accepting  $L' = L \cap H_g \in \text{Rat}(H_g)$  by Theorem 11. The labels of transitions are in  $H_g$ . We have  $g^{-1}H_gg \subseteq H$ . Hence it is enough to change every label  $h$  of transitions in  $\mathcal{A}'$  to  $g^{-1}hg$ . This gives the NFA  $\mathcal{A}$  for  $g^{-1}(L \cap H_g)g$  over  $H$ . ◀

**Proof of Theorem 23.** Let  $g \in G$  and  $K \in \text{Rat}(H)$ . First, we claim that we can rewrite  $Kg \in \text{Rat}(G)$  as a finite union of languages  $g'K'$  with  $g' \in G$  and  $K' \in \text{Rat}(H)$ .

Note that we can compute a set  $U_g \subseteq H$  of left-representatives such that  $H = \bigcup \{uH_g \mid u \in U_g\}$ . Indeed, by assumption, the membership to  $H$  is decidable, and hence the membership to  $gHg^{-1}$  and to  $H_g = gHg^{-1} \cap H$  is decidable, too. By the second assumption, we can compute the index  $k = |H : H_g|$ . Thus we can enumerate the elements of  $H$  until we find  $k$  elements that belong to  $k$  different left cosets of  $H_g$ . Checking if two elements belong to the same coset is decidable since the membership to  $H_g$  can be decided. Thus,

$$\begin{aligned} Kg &= \bigcup \{K \cap uH_g \mid u \in U_g\}g = \bigcup \{ugg^{-1}(u^{-1}K \cap H_g)g \mid u \in U_g\} \\ &= \bigcup \{g'g^{-1}(gg'^{-1}K \cap H_g)g \mid g' \in U_gg\}. \end{aligned}$$

Using Lemma 25 we obtain  $g^{-1}(gg'^{-1}K \cap H_g)g = K' \in \text{Rat}(H)$ . This shows the claim.

Let  $L \in \text{Frat}(G, H)$ . Hence,  $L$  is equal to a finite union of languages  $L_0g_1L_1 \cdots g_tL_t$  where all  $L_i \in \text{Rat}(H)$ . Using the claim, we can write  $L$  as a finite union of languages  $gK$  with  $g \in G$  and  $K \in \text{Rat}(H)$ . Since membership in  $H$  is decidable, we can effectively enumerate a set  $S$  of all distinct representatives of the right cosets of  $H$ , and moreover for each  $g \in G$  find a representative  $g' \in S$  such that  $g \in g'H$ . Since  $g = g'h$  for some  $h \in H$ , we can write  $gK = g'(hK)$ , where  $hK \in \text{Rat}(H)$ . Therefore, every flat rational set  $L$  can be written as a union  $L = \bigcup_{i=1}^n g_iK_i$ , where  $g_i \in S$  and  $K_i \in \text{Rat}(H)$ . Since  $gK_1 \cup gK_2 = g(K_1 \cup K_2)$ , we may assume that all  $g_i$  in the expression  $L = \bigcup_{i=1}^n g_iK_i$  are different.

Now let  $L$  and  $R$  be two flat rational sets. By the above argument we may assume that  $L = \bigcup_{i=1}^n a_iL_i$  and  $R = \bigcup_{j=1}^m b_jR_j$ , where  $a_i, b_j \in S$  and  $L_i, R_j \in \text{Rat}(H)$ . Then we have  $L \setminus R = \bigcup_{i=1}^n (a_iL_i \setminus \bigcup_{j=1}^m b_jR_j)$ . Note that if  $a_i \notin \{b_1, \dots, b_m\}$ , then  $a_iL_i \setminus \bigcup_{j=1}^m b_jR_j = a_iL_i$ , but if  $a_i = b_j$  for some  $j$  then  $a_iL_i \setminus \bigcup_{j=1}^m b_jR_j = a_i(L_i \setminus R_j)$ . Since  $\text{Rat}(H)$  is an effective relative Boolean algebra, we can compute the rational expression for  $L_i \setminus R_j$  in  $H$ . Hence we can compute the flat rational expression for  $L \setminus R$ . ◀

## 5.1 Reducing the membership problem of $\text{Frat}(M, G)$ to that of $\text{Frat}(M, H)$ for $|G/H| < \infty$

We suppose throughout this subsection that the group of units in  $M$  contains the subgroups  $G$  with a subgroup  $H$  such that the index of  $H$  in  $G$  is finite. It is clear that the membership problem of  $\text{Frat}(M, H)$  is a special case of the membership problem of  $\text{Frat}(M, G)$ . The aim is to prove the converse: the membership problem of  $\text{Frat}(M, G)$  is reducible to the membership problem of  $\text{Frat}(M, H)$ . The proof uses Theorem 11. In order to make the reduction effective, we impose some mild decidability conditions. As usual, elements of the monoid  $M$  are encoded by bit-strings. If the encoding of an element  $m \in M$  uses  $b$  bits, then we let  $\|m\|_{\text{bin}} = 1 + b$ . In particular, we can use (3) as the binary input size for an  $M$ -NFA  $\mathcal{A}$ . Next, we assume the following conditions.

1. The bit encoding is unique, thus  $m = m'$  if their bit encoding is the same.

2. Given the bit encodings of  $m, m' \in M$  we can compute the bit encoding of the product  $mm'$ .
3. The index  $|G/H|$  is known and given the bit encoding of an element  $m \in M$  we can decide whether  $m \in G$ ; and if yes, we can decide whether  $m \in H$ .

► **Theorem 26.** *Let  $G$  be a subgroup of the units in  $M$  and  $H \leq G$  be a finite index subgroup. Then membership problem of  $\text{Frat}(M, G)$  is reducible to the membership problem of  $\text{Frat}(M, H)$ .*

*More precisely, given  $m \in M$  and an  $M$ -NFA  $\mathcal{A}$  which is flat over  $G$ , then there is an  $M$ -NFA  $\mathcal{B}$  which is flat over  $H$  such that  $\|\mathcal{B}\|_{\text{bin}}$  is polynomial in  $\|\mathcal{A}\|_{\text{bin}}$  and we have  $m \in L(\mathcal{A}) \iff m \in L(\mathcal{B})$ . If  $M$  satisfies the mild decidability conditions above, then the construction is effective.*

The main ingredient of the following proof of Theorem 26 is the application of Theorem 11.

**Proof.** We assume that the input is specified by an  $M$ -NFA  $\mathcal{A} = (Q, \delta, q_{\text{in}}, q_{\text{fin}})$  which is flat over  $G$ , and by an element  $m \in M$ . W.l.o.g.,  $q_{\text{init}}$  is the unique initial state and without any incoming transition and  $q_{\text{fin}}$  the unique final state and without any outgoing transition. Moreover,  $q_{\text{in}} \neq q_{\text{fin}}$ . By adding, if necessary  $\varepsilon$ -selfloops, we may assume that all other states have incoming and outgoing transitions.

For  $i = 1, \dots, t$  let  $\mathcal{A}_i = (Q_i, \delta_i, I_i, F_i)$  be the set of (disjoint) subautomata of  $\mathcal{A}$  which are induced by the strongly connected components of  $\mathcal{A}$  with a nonempty set of transitions. Thus,  $q_{\text{in}}, q_{\text{fin}}$  are not included. Let  $1 \in R \subseteq G$  a finite set of right-coset representatives for  $H \backslash G$ . That is,  $G$  is the disjoint union  $G = \bigcup_{f \in R} Hf$  with  $1 \in R$ . For each  $1 \leq i \leq t$  and  $f \in R$  there is a  $G$ -NFA  $\mathcal{A}_{i,f} = (Q_{i,f}, \delta_{i,f}, I_{i,f}, F_{i,f})$  of polynomial size in  $\|\mathcal{A}\|_{\text{bin}}$  where  $Q_{i,f} = Q_i \times R$  such that  $L(\mathcal{A}_{i,f}) = L(\mathcal{A}_i) \cap Hf$ . Note that we have  $|\delta_{i,f}| \leq |\delta_i|$  by trimming. Hence,  $\sum_{1 \leq i \leq t} |\delta_{i,f}| \leq |\delta|$ . Moreover, w.l.o.g.  $I_{i,f} \leq I_i$  and  $F_{i,f} \leq F_i$ . The construction is effective, if  $M$  satisfies the mild decidability conditions above, then the construction is effective.

Introduce a new final state  $p_{i,f}$  and for each  $p \in F_{i,f}$  a new transition  $p \xrightarrow{\bar{f}} p_{i,f}$  where  $\bar{f} = f^{-1}$  in  $G$ . This leads to a new  $G$ -NFA  $\mathcal{A}'_{i,f} = (Q'_{i,f}, \delta'_{i,f}, I_{i,f}, \{p_{i,f}\})$  such that  $L(\mathcal{A}'_{i,f}) = L(\mathcal{A}_{i,f})f^{-1} \subseteq H$ . Since  $L(\mathcal{A}'_{i,f}) \in \text{Rat}(G)$ , we may apply Theorem 11. After renaming, we obtain an  $H$ -NFA  $\mathcal{B}_{i,f} = (Q'_{i,f}, \delta''_{i,f}, I_{i,f}, \{p_{i,f}\})$  such that  $L(\mathcal{B}_{i,f}) = L(\mathcal{A}'_{i,f})$ .

We are almost done. We begin with a disjoint union

$$\mathcal{B} = \{q_{\text{in}}, q_{\text{fin}}\} \cup \bigcup_{1 \leq i \leq t, f \in R} \mathcal{B}_{i,f} \quad (4)$$

Thus,  $q_{\text{in}}$  and  $q_{\text{fin}}$  are reintroduced for the same purpose:  $q_{\text{in}}$  becomes the unique initial state and  $q_{\text{fin}}$  becomes the unique final state.

For all  $f \in F$  we let  $Q_{0,f} = \{q_{\text{in}}\}$  and  $Q_{t+1,f} = \{q_{\text{fin}}\}$ . One after another consider all pairs  $(i, j)$  where  $0 \leq i, j \leq t+1$  and  $i \neq j$ . Then introduce for every transition  $p_i \xrightarrow{m_{i,j}} q_j \in \delta$  with  $p_i \in Q_i$  and  $q_j \in Q_j$  for every  $f \in R$  a new transition  $p_{i,f} \xrightarrow{f m_{i,j}} q'_{j,f} \in \delta$  where  $p_{i,f}$  is the unique final state in  $\mathcal{B}_{i,f}$  and  $q'_{j,f}$  is an initial state in  $\mathcal{B}_{j,f}$ . ◀

► **Example 27.** The construction in Theorem 26 is, in particular, effective in the following setting. Let  $\mathcal{A}$  be  $\text{GL}(2, \mathbb{Z})$ -NFA with  $n$  states which is flat over  $\text{SL}(2, \mathbb{Z})$ . Then, on input  $\mathcal{A}$ , we can compute in deterministic polynomial time an  $\text{SL}(2, \mathbb{Z})$ -NFA  $\mathcal{A}'$  with at most  $n$  states such that  $1 \in L(\mathcal{A}) \iff 1 \in L(\mathcal{A}')$ .

To see this, let us recall the construction underlying the proof of Theorem 11 which is the core for the proof of Theorem 26. For that we first define the matrix  $s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Hence,  $\text{GL}(2, \mathbb{Z}) = \text{SL}(2, \mathbb{Z}) \cup \text{SL}(2, \mathbb{Z})s$  is a disjoint union and  $s^2 = 1$ . Next, let  $Q'$  be the disjoint copy of the state space  $Q$  of  $\mathcal{A}$ . For a state  $p \in Q$  we denote by  $p'$  the corresponding copy in  $Q'$ . Without restriction we may assume that  $\mathcal{A}$  is trim. We let  $\mathcal{B} = Q \cup Q'$  a new state space and for each transition  $p \xrightarrow{a} q$  in  $\mathcal{A}$  we introduce new transitions, depending on whether  $a \in \text{SL}(2, \mathbb{Z})$ . For  $a \in \text{SL}(2, \mathbb{Z})$  we use the left picture and for  $a \in \text{GL}(2, \mathbb{Z}) \setminus \text{SL}(2, \mathbb{Z})$  we use the right one. (Note that the positions of  $q$  and  $q'$  interchanged in order to avoid to draw crossing edges.)





Clearly,  $1 \in L(\mathcal{A}) \iff 1 \in L(\mathcal{B})$ . In order to define  $\mathcal{A}'$  we keep the initial and final states of  $\mathcal{A}$ , but we remove all transitions which do not have a label in  $\text{SL}(2, \mathbb{Z})$ . This leads to an  $\text{SL}(2, \mathbb{Z})$ -NFA  $\mathcal{B}'$  with  $1 \in L(\mathcal{A}) \iff 1 \in L(\mathcal{B}')$ . Finally, we trim the NFA  $\mathcal{B}'$ ; and that yields the desired NFA  $\mathcal{A}'$ . Note that in the trimming process for each  $p \in Q$  exactly one of the states  $p, p'$  survives. Moreover, the initial and final of  $\mathcal{A}$  survive since we assumed  $\mathcal{A}$  to be trim.

► **Corollary 28.** *Let  $T$  be any subset of  $\mathbb{Q}^{2 \times 2}$ . Then the following task can be computed in deterministic polynomial time.*

*The input is a  $T$ -NFA  $\mathcal{A}$  which is flat over  $\text{GL}(2, \mathbb{Z})$ . The output is a  $T$ -NFA  $\mathcal{A}'$  which is flat over  $\text{SL}(2, \mathbb{Z})$  such that*

$$L(\mathcal{A}) = L(\mathcal{A}') \tag{5}$$

*That is: for  $2 \times 2$  matrices with entries in rational numbers there is a  $\mathbf{P}$ -reduction of the membership problem for flat rational sets over  $\text{GL}(2, \mathbb{Z})$  to that over  $\text{SL}(2, \mathbb{Z})$ .*

**Proof.** We may assume that  $\mathcal{A}$  is trim. First we compute the  $\{\mathcal{A}_i \mid 1 \leq i \leq k\}$  of strongly connected components of  $\mathcal{A}$ . By modifying  $\mathcal{A}$  we may assume that for entering each  $\mathcal{A}_i$  one has to use a unique  $\varepsilon$ -transition  $\tilde{p}_i \xrightarrow{1} p_i$  and in order to leave it, one has to use a unique  $\varepsilon$ -transition  $q_i \xrightarrow{1} \tilde{q}_i$  where  $p_i \neq q_i$ . Moreover, no state in any  $\mathcal{A}_i$  is initial or final with respect to the NFA  $\mathcal{A}$ . We view each  $\mathcal{A}_i$  as a  $\text{GL}(2, \mathbb{Z})$ -NFA. Let  $\mathcal{A}'_i$  be a disjoint copy of  $\mathcal{A}_i$ . By a slight modification of  $\mathcal{A}_i$  and  $\mathcal{A}'_i$  we may assume that  $p_i$  is the unique initial state of  $\mathcal{A}_i$  without incoming transitions and that  $p'_i$  is the unique initial state of  $\mathcal{A}'_i$  without incoming transitions. Using essentially the very same construction as in Example 27 (which was illustrating a special case for the use of Theorem 26) we can make another modification to  $\mathcal{A}_i$  and  $\mathcal{A}'_i$  (without using additional states) such that first,  $L(\mathcal{A}_i) = L(\mathcal{A}_i) \cap \text{SL}(2, \mathbb{Z})$  and  $L(\mathcal{A}'_i) = sL(\mathcal{A}_i) \cap \text{SL}(2, \mathbb{Z})$  where  $s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ , and second,  $\mathcal{A}_i$  and  $\mathcal{A}'_i$  are both  $\text{SL}(2, \mathbb{Z})$ -NFAs. Finally, we connect each  $\mathcal{A}'_i$  to the ambient NFA  $\mathcal{A}$  by a transition  $\tilde{p}_i \xrightarrow{s} p_i$  (which has label  $s$ ) and an  $\varepsilon$ -transition  $q'_i \xrightarrow{1} \tilde{q}_i$ . This yields a  $T$ -NFA  $\mathcal{A}'$  being flat over  $\text{SL}(2, \mathbb{Z})$  such that  $L(\mathcal{A}) = L(\mathcal{A}')$ . ◀

Corollary 28 implies that it does not really matter if we consider  $2 \times 2$  matrices with respect to flatness over  $\text{SL}(2, \mathbb{Z})$  or  $\text{GL}(2, \mathbb{Z})$ . While flatness over  $\text{GL}(2, \mathbb{Z})$  yields, formally, stronger statements, the use of  $\text{SL}(2, \mathbb{Z})$  is sometimes more convenient.

## 6 Membership to flat rational sets of nonsingular matrices

The aim of this section is to investigate the time complexity of deciding the membership problem in  $\text{Frat}(\text{GL}(2, \mathbb{Q}), \text{GL}(2, \mathbb{Z}))$ . In Section 6.3 we generalize the result to  $\text{Frat}(\text{GL}(2, \mathbb{Q}), M)$ , where  $M$  is the submonoid of  $\text{GL}(2, \mathbb{Q})$  that is equal to the union of  $\text{GL}(2, \mathbb{Z})$  and the set of matrices  $g \in \text{GL}(2, \mathbb{Q})$  with  $|\det(g)| > 1$ .

In the following, whenever we consider a class  $\text{Frat}(\mathbb{Q}^{2 \times 2}, N)$ , then the notation NFA always refers to an  $\mathbb{Q}^{2 \times 2}$ -NFA. Thus, the transitions are labeled with  $2 \times 2$  matrices over  $\mathbb{Q}$ . If an NFA only uses transitions in  $\mathbb{Z}^{2 \times 2}$ , then we call the NFA an  $\mathbb{Z}^{2 \times 2}$ -NFA.

We are interested in the time complexity of the following decision problem “ $g \in L(\mathcal{A})$ ?” where NFA  $\mathcal{A}$  is flat over  $\text{GL}(2, \mathbb{Z})$ . For the input size we use binary encoding. Recall that the binary input size of a matrix and of an NFA was defined in Section 3.1. Since we can construct in polynomial time an NFA which is flat over  $\text{GL}(2, \mathbb{Z})$  and which accepts  $g^{-1}L(\mathcal{A})$ , it is enough to consider the following special case.

▷ Problem 1.

INPUT: A  $\mathbb{Q}^{2 \times 2}$ -NFA  $\mathcal{A}$  which is flat over  $\text{GL}(2, \mathbb{Z})$ .

QUESTION: “ $1 \in L(\mathcal{A})$ ?”

## 6.1 Membership for $\text{SL}(2, \mathbb{Z})$

Let us recall the simplest version Problem 1. We begin with an NFA  $\mathcal{A}$  such that all matrices are in  $\text{SL}(2, \mathbb{Z})$  and we ask “ $1 \in L(\mathcal{A})$ ?”. So far it is known that the problem is **NP**-hard and stated to be **NP**-complete for by  $\text{PSL}(2, \mathbb{Z})$ [7]. For the subgroup membership problem better complexity bounds are known due to Lohrey, [28]: If  $L(\mathcal{A})$  defines a subgroup of  $\text{GL}(2, \mathbb{Z})$ , then, on input  $\mathcal{A}$  where matrix entries are written in binary, the question “ $1 \in L(\mathcal{A})$ ?” can be answered in **P**. This generalizes a result of [21] for the modular group  $\text{PSL}(2, \mathbb{Z})$ . Actually, Lohrey shows an even stronger result using the notion of *power words*. However, it is not known how to apply these **P**-results to the membership problem for flat rational sets. We therefore content ourselves with the following pseudo-polynomial time result. (The complexity is *pseudo-polynomial* if it is polynomial time when integers are given in unary.)

► **Proposition 29.** *There is some fixed constant  $\kappa$  such that the following decision problem can be decided in  $\text{DTIME}((sM)^\kappa)$ .*

INPUT: An NFA  $\mathcal{A}$  with  $s$  transitions and with labels in  $\text{SL}(2, \mathbb{Z})$  and where  $M$  is the maximal absolute value of an integer entry appearing in matrix which labels a transition.

QUESTION: “ $1 \in L(\mathcal{A})$ ?”

**Proof.** Since the factor commutator subgroup of  $\text{SL}(2, \mathbb{Z})$  is a free subgroup of rank 2 and of index 12 (by [35]), we can reduce the problem in polynomial time to the special instance where all matrices are in a free group  $F$  with a finite generating set  $\Sigma = \Sigma^{-1}$ . In a second polynomial time reduction we replace matrices by words over  $\Sigma$ . More precisely, based on [21] (actually a refinement in [14]) it is possible to replace the matrices in  $F$  by words over  $\Sigma$  where each transition is labeled with word of length in  $\mathcal{O}(1)$  such that the number of transitions is in  $\mathcal{O}(sM)$ . Formally, this is now an NFA  $\mathcal{A}'$  with  $L(\mathcal{A}') \subseteq \Sigma^*$ . The details are in [13]. Having this we apply Benois’ result on free groups, [8]. The result is another automaton  $L(\mathcal{B}')$  having, up to a constant factor, the same number of states as  $\mathcal{B}$  such that  $\mathcal{B}'$  accepts as an NFA over  $\Sigma^*$  exactly the reduced normal forms belonging to  $L(\mathcal{A})$ . The answer to the question “ $1 \in L(\mathcal{A})$ ?” is now the same as the answer to question whether the empty word is accepted by  $\mathcal{B}'$ . This can be done in polynomial time with respect to  $sM$ . ◀

► **Remark 30.** If we started with an input where matrices are written in binary, then the above statement shows decidability in **DEXPTIME** by using Benois’ algorithm in [8]. Possibly, a more sophisticated method could lead to solving Problem 1 in **NP** for binary inputs.

## 6.2 Membership for $\text{Frat}(\mathbb{Q}^{2 \times 2}, \text{GL}(2, \mathbb{Z}))$ for non-singular matrices

Throughout, an  $\varepsilon$ -transition means a transition  $p \xrightarrow{1} q$  where  $1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  denotes the identity matrix.

► **Lemma 31.** *There is an **NP** reduction of Problem 1 to following problem.*

▷ Problem 2. The input is written in binary.

INPUT: A positive integer  $m$  and an  $\mathbb{Z}^{2 \times 2}$ -NFA  $\mathcal{A}$  which is flat over  $\text{GL}(2, \mathbb{Z})$ .

QUESTION: “ $m \in L(\mathcal{A})$ ?”

Moreover, if  $D$  denotes the maximal absolute value  $|\det(p_m m)|$  over all labels  $m \in \text{GL}(2, \mathbb{Q})$  which appear on transitions of  $\mathcal{A}$  and where  $p_m$  is the least positive integer such that  $p_m m \in \mathbb{Z}^{2 \times 2}$ , then **NP** reduction does not increase value  $D$ .

**Proof.** Using Corollary 28 and the fact that the input 1 to Problem 1 is nonsingular, it is enough to consider the case where  $\mathcal{A}$  is an  $\text{GL}(2, \mathbb{Q})$ -NFA being flat over  $\text{SL}(2, \mathbb{Z})$ . If  $1 \in L(\mathcal{A})$ , then there is an accepting path where transitions outside  $\text{SL}(2, \mathbb{Z})$  are used  $t$  times where  $t$  is less than the number of strongly connected components of  $\mathcal{A}$ . Thus, we can guess an initial state  $q_0$  and a sequence of  $t$  transitions  $q_{j-1} \xrightarrow{g_j} q_j$  for  $1 \leq j \leq t$  such that all other transitions (which are used on that path) are labeled with matrices from  $\text{SL}(2, \mathbb{Z})$ . W.l.o.g.  $t \geq 1$  since we may assume that no initial state in  $\mathcal{A}$  is final.

Thus, by making the NFA not larger, having done the guesses above, we compute in polynomial time  $t$  subautomata  $\mathcal{A}_i$  of  $\mathcal{A}$  for  $1 \leq i \leq t$  such that  $1 \in L(\mathcal{A})$  implies

$$1 \in g_1 L(\mathcal{A}_1) g_2 L(\mathcal{A}_2) \cdots g_t L(\mathcal{A}_t). \quad (6)$$

Let  $R_j = L(\mathcal{A}_j)$ . Now for each  $1 \leq j \leq t$  let  $p_j$  be the least positive integer that  $p_j g_j \in \mathbb{Z}^{2 \times 2}$  and  $m = \prod_j p_j$ . Let  $g'_j = p_j g_j$ . Suppose that  $\|g_j\|_{\text{bin}} \leq n$  for all  $j$ , then Lemma 14 implies that  $\prod_j \|g'_j\|_{\text{bin}} \in \mathcal{O}(n^2)$  which is polynomial and therefore acceptable. Next, we have

$$1 \in g_1 R_1 \cdots g_t R_t \iff m \in g'_1 R_1 \cdots g'_t R_t.$$

Thanks to the definition of the  $g'_j$ 's, the value  $D$  is not increasing. ◀

By Lemma 31 we content ourselves to study the membership in  $\text{Frat}(\text{GL}(2, \mathbb{Q}), \text{SL}(2, \mathbb{Z}))$  in the special case where all matrices involved are integer matrices.

► **Definition 32.** Let  $0 \neq q \in \mathbb{Z}$  and  $e \in \{b, c\}$ . Then we denote

$$H_{e,q} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z}) \mid e \equiv 0 \pmod{q} \right\}.$$

► **Lemma 33.** The subgroups  $H_{b,q}$  and  $H_{c,q}$  of  $\text{SL}(2, \mathbb{Z})$  in Definition 32 are conjugated:

$$\begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}^{-1} H_{c,q} \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} = H_{b,q}. \quad (7)$$

Moreover their index in  $\text{SL}(2, \mathbb{Z})$  is less than  $q^2$ . In particular, they are of finite index and therefore recognizable subsets in  $\text{SL}(2, \mathbb{Z})$  and  $\text{GL}(2, \mathbb{Z})$ .

**Proof.** Calculating mod  $q$  we see that  $H_{b,q}$  and  $H_{c,q}$  are subgroups of the normal subgroup where the determinant is 1 mod  $q$ . This normal subgroup has an index about  $q^3$ . However, we can be more precise. The image of  $H_{b,q}$  in  $\text{SL}(2, \mathbb{Z}/q\mathbb{Z})$  is the group

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z}/q\mathbb{Z}) \mid b = 0 \right\}.$$

Thus, the image has size  $|(\mathbb{Z}/q\mathbb{Z})^* \times \mathbb{Z}/q\mathbb{Z}|$  which is at most  $(q-1)q < q^2$ . ◀

► **Lemma 34.** Let  $q \in \mathbb{Z}$  and  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$  be given in binary encoding. Then for  $e \in \{b, c\}$  there is a matrix  $m = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$  where  $\|m\|_{\text{max}} \leq q$  such that  $g \in m H_{e,q}$ . In particular,  $\|m\|_{\text{bin}} \leq \log(q)$ ; and we can guess the matrix  $m$  in **NP** and verify in polynomial time that  $m^{-1}g \in H_{e,q}$ .

**Proof.** Since  $g \in \text{SL}(2, \mathbb{Z})$ , the entries  $b$  and  $d$  are coprime. By Lemma 1 there are coprime  $x$  and  $y$  such that first,  $xb + yd \equiv 0 \pmod{q}$  and second  $1 \leq x, y \leq |q|$ . We can guess  $x, y$  within **NP**. Next, we apply Proposition 2 to obtain  $w, z$  in polynomial time such that first,  $xz - yw = 1$  and second  $1 \leq |w|, |z| \leq \max\{x, y\} \leq |q|$ . We obtain  $\begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in H_{b,q}$ . This shows the result for  $e = b$ . The result for  $e = c$  is symmetric. ◀

► **Theorem 35.** Let  $g \in \text{GL}(2, \mathbb{Q})$  and  $R = L(\mathcal{A})$  such that the NFA  $\mathcal{A}$  is flat over  $\text{GL}(2, \mathbb{Z})$  and the matrix  $g$  and all matrices which appear as labels of transitions are nonsingular matrices with entries in  $\mathbb{Q}$ . Let  $D$  denote the maximal absolute value  $|\det(p_h h)|$  over all labels  $h \in \text{GL}(2, \mathbb{Q})$  which

appear on transitions of  $\mathcal{A}$  and where each  $p_h$  is the least positive integer such that  $p_h h \in \mathbb{Z}^{2 \times 2}$ . Finally, let  $n = \|g\|_{\text{bin}} + \|\mathcal{A}\|_{\text{bin}}$  be the input size. Then the problem “ $g \in R?$ ” can be decided in  $\text{DTIME}(2^{n^{\mathcal{O}(1)}}) = \text{DEXPTIME}$  and, moreover, there is a  $\text{NTIME}(D^{\mathcal{O}(n)} n^{\mathcal{O}(1)})$  reduction of the decision problem “ $g \in R?$ ” to a problem  $1 \in R'$  where  $R' \in \text{RAT}(\text{SL}(2, \mathbb{Z}))$  and  $R'$  is given by some NFA  $\mathcal{A}'$  where all labels of transitions are matrices in  $\text{SL}(2, \mathbb{Z})$ .

**Proof.** We begin by showing that there exists a constant  $\kappa$  such that there exists the claimed  $\text{NTIME}(D^{\mathcal{O}(n)} n^\kappa)$  reduction. In the first step, we apply the **NP**-reduction in Lemma 31. Recall that the reduction does not increase  $D$ . Thus, it suffices to deal with Problem 2. This means the starting point is a  $\mathbb{Z}^{2 \times 2} \cap \text{GL}(2, \mathbb{Q})$ -NFA  $\mathcal{A}$  which is flat over  $\text{SL}(2, \mathbb{Z})$  and the question is  $m \in L(\mathcal{A})$  for some positive integer  $m$ . Hence, we assume that  $g$  and all labels  $h \in \text{GL}(2, \mathbb{Q})$  which appear on transitions of  $\mathcal{A}$  are integer matrices. In particular,  $p_h = 1$  for all matrices  $h$  and  $D$  is the maximal absolute value  $|\det(h)|$  over all labels  $h$  which appear in  $\mathcal{A}$ . In case that  $D = 1$  we are done with the reduction since  $L(\mathcal{A}) \subseteq \text{GL}(2, \mathbb{Z})$ . Since  $g \in \text{GL}(2, \mathbb{Z})$  we may assume without restriction that  $g = 1$ . Thus, the question is whether  $1 \in L(\mathcal{A})$  where  $\mathcal{A}$  is a  $\text{GL}(2, \mathbb{Z})$ -NFA. We are done by Corollary 28, since we can transform  $\mathcal{A}$  in deterministic polynomial time into an  $\text{SL}(2, \mathbb{Z})$ -NFA  $\mathcal{A}'$  such that  $1 \in L(\mathcal{A}) \iff 1 \in L(\mathcal{A}')$ . We assume that  $g$  and all labels  $h \in \text{GL}(2, \mathbb{Q})$  which appear on transitions of  $\mathcal{A}$  are integer matrices. In particular,  $p_h = 1$  for all matrices  $h$  and  $D$  is the maximal absolute value  $|\det(h)|$  over all labels  $h$  which appear in  $\mathcal{A}$ . In case that  $D = 1$  we are done with the reduction since  $L(\mathcal{A}) \subseteq \text{GL}(2, \mathbb{Z})$  and then we can apply Theorem 26.

Thus, henceforth, we may assume without restriction that  $D \geq 2$ . Since  $\mathcal{A}$  is flat over  $\text{GL}(2, \mathbb{Z})$ , we can guess in a second step within **NP** the following items: a natural number  $t \leq n$ , matrices  $g_1, \dots, g_t$  which appear as labels of transitions in  $\mathcal{A}$ , and  $t$  (disjoint) trim subautomata  $\mathcal{A}_i$  of  $\mathcal{A}$  such that

$$g \in R \iff \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \in g_1 R_1 \cdots g_t R_t. \quad (8)$$

In (8) we have  $0 < m \in \mathbb{N}$  and  $R_i = L(\mathcal{A}_i) \subseteq \text{SL}(2, \mathbb{Z})$  for  $1 \leq i \leq t \leq n$ . We perform the following steps.

1. For each  $i$  with the help of two additional states  $p_i$  and  $q_i$  two  $\varepsilon$ -transitions  $p_i \xrightarrow{\varepsilon} p'_i$  and  $q'_i \xrightarrow{\varepsilon} q_i$  we have the following: Each  $R_i$  is given by some NFA with a single initial state  $p_i$  without incoming edge and a single final state  $q_i$  without outgoing edge. Moreover,  $p_i$  has exactly one outgoing transition  $p_i \xrightarrow{m_i} p'_i$  and  $q_i$  has exactly one incoming transition  $q'_i \xrightarrow{n_i} q_i$  where  $m_i, n_i \in \text{SL}(2, \mathbb{Z})$ .

We keep this property as a loop-invariant. In the beginning we let  $m_i = n_i = 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ .

2. For each  $1 \leq i \leq t$  compute for  $g_i$  its Smith normal form  $g_i = r_i e_i \begin{pmatrix} 1 & 0 \\ 0 & q_i \end{pmatrix} f_i$  in polynomial time. Replace each  $g_i$  by  $g'_i = e_i \begin{pmatrix} 1 & 0 \\ 0 & q_i \end{pmatrix} f_i$  and the left hand side  $r$  by  $r' = r / \prod_i r_i$ . We then ask whether  $r' \in g'_1 R_1 \cdots g'_t R_t$ . By conjugation with  $e_1$ , question is equivalent to the question

$$“r' \in \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} f_1 R_1 g'_2 R_2 \cdots g'_t R_t e_1?”.$$

Note that the product  $\prod_{1 \leq i \leq t} r_i^2 q_i \leq D^t < D^n$ . Writing down the product  $r = \prod_{1 \leq i \leq t} r_i$  in binary requires at most  $n \log(D)/2$  bits.

3. For  $1 \leq i \leq t$ , replace the transitions  $p_i \xrightarrow{m_i} p'_i$  by  $p_i \xrightarrow{f_i \cdot m_i} p'_i$ .
4. For  $1 \leq i \leq t$ , replace the transitions  $q'_i \xrightarrow{m_i} q_i$  by  $q'_i \xrightarrow{m_i \cdot e_{i+1}} q_i$  where we let  $e_{t+1} = e_1$ . These additional transitions replace  $R_i$  with  $R'_i = f_i R_i e_{i+1}$  for  $1 \leq i \leq t$ , where  $e_{t+1} = e_1$ . To simplify the notation, we rename  $R'_i$  back to  $R_i$ .
5. Adding less than  $3t$   $\varepsilon$ -transitions, we may assume without restriction that  $t$  is a power of 2. Thus,  $t \in 2^{\mathbb{N}}$ .

The last step finishes the preprocessing phase. After that the problem is to decide whether

$$r \in g_1 R_1 \cdots g_t R_t \quad (9)$$

where  $r \in \mathbb{Q}$  is positive,  $R_i \in \text{Rat}(\text{SL}(2, \mathbb{Z}))$ , and  $g_i = \begin{pmatrix} 1 & 0 \\ 0 & q_i \end{pmatrix}$  for all  $1 \leq i \leq t \in 2^{\mathbb{N}}$  with  $0 \neq q_i \in \mathbb{Z}$ . Each  $R_i$  is represented by some NFA  $\mathcal{A}_i$  such that  $R_i = L(\mathcal{A}_i)$ . Let  $q = \max \{|q_i| \mid 1 \leq i \leq t\}$ .

Then we have  $q \leq D$  we aim now at a  $\text{NTIME}(q^{\mathcal{O}(n)} n^\kappa)$  reduction. This suffices and it facilitates the following explanation. Without restriction we assume  $t \geq 2$ . We perform  $\log t$  rounds. In round  $k = 1, \dots, \log t$  we let  $s = t/2^{k-1}$ . Each round  $s$  starts with a question “ $r_s \in g_{1,s} R_{1,s} \cdots g_{t_s,s} R_{t_s,s}$ ?” where  $0 < r_s \in \mathbb{Q}$ ,  $R_{i,s} = L(\mathcal{A}_{i,s}) \in \text{Rat}(\text{SL}(2, \mathbb{Z}))$ , and  $g_{i,s} = \begin{pmatrix} 1 & 0 \\ 0 & q_{i,s} \end{pmatrix}$  for all  $i, s$  with  $0 \neq q_{i,s} \in \mathbb{Z}$ . Moreover,  $t_s \in 2^{\mathbb{N}}$ . The first round starts with “ $r \in g_1 R_1 \cdots g_t R_t$ ?” which was defined in (8) and which has the form above. After that each round will halve the number  $s$  until  $s$  becomes 1 where we stop. In the  $k$ -th round we perform the following steps where  $s = t/2^{k-1}$ .

1. For sake of a simplified notation we rename  $R_{i,s}$  and  $L(\mathcal{A}_{i,s})$  as  $R_i$  resp. as  $L(\mathcal{A}_i)$ . Thus, the question in the  $k$ -th round becomes the question

$$“r \in \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} R_1 \cdots \begin{pmatrix} 1 & 0 \\ 0 & q_s \end{pmatrix} R_s?”$$

2. By Lemma 33 we can write  $\text{SL}(2, \mathbb{Z}) = \bigcup_{i \in I} \{m H_{c,q_i} \mid m \in S_{q_i}\}$  where  $|I| \leq q_i^2$ . Moreover, by Lemma 34 we can choose finite sets  $S_{q_i} \subseteq \text{SL}(2, \mathbb{Z})$  such  $\|m\|_{\text{bin}} \leq \log q_i$  for all  $m \in S_{q_i}$ . Therefore, the “witness” for  $r \in \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} R_1 \cdots \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} R_s$  is a sequence of  $s - 1$  matrices  $(m_2, \dots, m_s)$  in  $\text{SL}(2, \mathbb{Z})$  such that

$$r \in \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} (R_1 \cap m_2 H_{c,q_2}) \cdots \begin{pmatrix} 1 & 0 \\ 0 & q_{s-1} \end{pmatrix} (R_{s-1} \cap m_s H_{c,q_s}) \begin{pmatrix} 1 & 0 \\ 0 & q_s \end{pmatrix} R_s \text{ with } m_j \in S_{q_j}. \quad (10)$$

Hence,  $\|m_j\|_{\text{bin}} \leq \log(q_j)$ .

3. With a new single initial state and additional transitions (and after a renaming) we may assume that  $L(\mathcal{A}_i) = m_i^{-1} R_{i-1}$  for all  $i \geq 2$ . We don't touch the  $L(\mathcal{A}_1)$ .
4. Again, using Lemma 33, we know that the subgroup  $H_{c,q_i}$  is of finite index less than  $q_i^2$  in  $\text{SL}(2, \mathbb{Z})$ ; and thus,  $H_{c,q_i}$  is recognizable. This implies  $L_i = m_i^{-1} R_{i-1} \cap H_{c,q_i}$  is rational in  $\text{SL}(2, \mathbb{Z})$ . (As stated in the proof of Lemma 9, the implication is a general fact for monoids.) More precisely, for all even  $i$ , starting with the product NFA of  $\mathcal{A}_i$  with an NFA having as states the right cosets  $H_{c,q_i} \backslash \text{SL}(2, \mathbb{Z})$ , we construct another NFA  $\mathcal{B}_i$  such that first, the NFA accepts  $L_i = m_i^{-1} R_{i-1} \cap H_{c,q_i}$  and second, all labels of transitions are in  $H_{c,q_i}$ . By the techniques used in the proof of Theorem 11 the construction can be done in time which is a polynomial in  $q_i$ .
5. For all  $i$  redefine  $g_i = \begin{pmatrix} 1 & 0 \\ 0 & q_i \end{pmatrix}$ . For every even  $i$  we write

$$\begin{aligned} (R_{i-1} \cap m_i H_{c,q_i}) g_i &= m_i (m_i^{-1} R_{i-1} \cap H_{c,q_i}) g_i \\ &= m_i g_i (g_i^{-1} (m_i^{-1} R_{i-1} \cap H_{c,q_i}) g_i) = m_i g_i (g_i^{-1} L_j g_i). \end{aligned}$$

6. Define  $K_i = \bar{g}_i L_i g_i$ . The NFA for accepting  $K_i$  is the NFA  $\mathcal{B}_i$  where every label  $m \in H_{c,q_i}$  is replaced by  $\bar{g}_i m g_i$ . Thus, the new labels belong to the subgroup  $H_{b,q_i}$  of  $\text{SL}(2, \mathbb{Z})$ .
7. Define  $R'_i = K_i \cdot R_i$  and let  $g'_i = g_{i-1} m_i g_i$ . Compute the Smith normal form of  $g'_i = r'_i e'_i \begin{pmatrix} 1 & 0 \\ 0 & q'_i \end{pmatrix} f'_i$  in time which is fixed polynomial in  $\log(q_i)$ . This follows from Lemma 34. Thus, for some fixed constant  $\kappa$  we have  $|q'_i| \leq |q_i^\kappa|$  and therefore the binary presentation of  $q'_i$  is linear in the binary presentation of  $q_i$ .
8. Repeat what we have done above. First, for each  $i$  push the positive  $r'_i$  integers to left by multiplying the left side of the question with  $1/r'_i$ . This yields a new positive rational number  $r'$  on the left side. In addition, the  $g'_i$  become equal to  $e'_i \begin{pmatrix} 1 & 0 \\ 0 & q'_i \end{pmatrix} f'_i$ . Second, replace  $R'_i$  by  $R''_i = f'_i R'_i e'_{i+1}$  where  $e'_{s+1} = e'_s$ . Third, let  $e''_i = \begin{pmatrix} 1 & 0 \\ 0 & q'_i \end{pmatrix}$ . Overall, we ask now the question “ $r' \in \begin{pmatrix} 1 & 0 \\ 0 & q'_1 \end{pmatrix} R''_1 \cdots \begin{pmatrix} 1 & 0 \\ 0 & q'_{s/2} \end{pmatrix} R''_{s/2}$ ?”. If  $s/2 + 1$ , then we have  $q'_1 = 1$ .

This finishes the non-deterministic reduction after the preprocessing. For the time complexity observe that after  $\log t$  rounds the largest value  $|q_i|$  is bounded by  $q^{\mathcal{O}(\log n)} \leq 2^{\mathcal{O}(n \log n)} \leq 2^{\mathcal{O}(n^2)}$ . However, due to the product automaton construction we need for some fixed  $\kappa$  nondeterministic time  $\text{NTIME}(q^{\mathcal{O}(n)} n^\kappa)$ . Together with the preprocessing we obtain  $\text{NTIME}(D^{\mathcal{O}(n)} n^\kappa)$ .

If we avoid all nondeterministic guesses of matrices, then we can run a deterministic reduction in  $\text{DTIME}(2^{n^{\mathcal{O}(1)}})$ . Finally, we apply Proposition 29. It uses two parameters  $s$  and  $M$ . Here,  $s$  is the number of transitions and  $M$  is the maximal  $\|m\|$  where  $m$  runs over the label of transitions. The construction above shows  $s \in \text{DTIME}(2^{n^{\mathcal{O}(1)}})$  and  $M \in \text{DTIME}(2^{n^{\mathcal{O}(1)}})$ . Therefore,  $sM \in \text{DTIME}(2^{n^{\mathcal{O}(1)}})$  and we are done.  $\blacktriangleleft$

### 6.3 Flat over $g \in \text{GL}(2, \mathbb{Z}) \cup \{g \in \text{GL}(2, \mathbb{Q}) \mid |\det(g)| > 1\}$

So far, we considered  $\text{Frat}(\text{GL}(2, \mathbb{Q}), M)$  only when  $M = \text{GL}(2, \mathbb{Z})$  or  $M = \text{SL}(2, \mathbb{Z})$ . In this subsection we show that membership to  $\text{Frat}(\text{GL}(2, \mathbb{Q}), M)$  is decidable for the monoid

$$M = \text{GL}(2, \mathbb{Z}) \cup \{g \in \text{GL}(2, \mathbb{Q}) \mid |\det(g)| > 1\}. \quad (11)$$

Theorem 36 generalizes Theorem 35 with respect to decidability. However, not very surprisingly, our complexity estimation becomes worse. We consider the following problem

▷ **Problem 3.**

INPUT: An  $\text{GL}(2, \mathbb{Q})$ -NFA  $\mathcal{A}$  which is flat over  $\text{GL}(2, \mathbb{Z}) \cup \{g \in \text{GL}(2, \mathbb{Q}) \mid |\det(g)| > 1\}$  where matrices are written in binary.

QUESTION: “ $1 \in L(\mathcal{A})$ ?”

► **Theorem 36.** *Problem 3 is decidable in deterministic doubly exponential time  $\text{DTIME}(2^{2^p})$ , where  $p$  is some polynomial.*

**Proof.** As we did in the proof of Problem 1, it no restriction to start in Problem 3 with a question

$$“r \in g_1 L(\mathcal{A}_1) g_2 L(\mathcal{A}_2) \cdots g_t L(\mathcal{A}_t) ?” \quad (12)$$

Here  $r \in \mathbb{Q}$  is positive and we have  $g_i \in \mathbb{Z}^{2 \times 2} \cap \text{GL}(2, \mathbb{Q})$  and  $\mathcal{A}_i$  is an  $M$ -NFA for all  $1 \leq i \leq t$ . Thus, each  $g_i$  belongs to  $M$ , and we can work with a single  $M$ -NFA  $\mathcal{A}'$  and we ask “ $r \in L(\mathcal{A}')$ ?”. Let us define  $n = \left\| \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \right\|_{\text{bin}} + \|\mathcal{A}\|_{\text{bin}}$ , then  $n$  is polynomial in the original binary input size; which we can ignore. The new starting point is the question “ $r \in L(\mathcal{A}')$ ?”. For simplicity of notation, we rename  $\mathcal{A}'$  as  $\mathcal{A}$ .

Let  $T = \{g \in \text{GL}(2, \mathbb{Q}) \mid |\det(g)| > 1 \text{ and } g \text{ is the label of a transition in } \mathcal{A}\}$ . If  $T = \emptyset$  we are done. Hence, we may assume  $T \neq \emptyset$ . Let  $N \in \mathbb{N}$  be the smallest number such that  $1 + 1/N \leq |\det(g)|$  for all  $g \in T$ . A direct calculation shows  $N \leq 2^{2n+1}$ . If  $r \in L(\mathcal{A})$ , then there is a minimal  $k \in \mathbb{N}$  such that transitions with label in  $T$  are used exactly  $k$  times. Let give an upper bound of  $k$ . For  $N = 1$  we have  $2^k \leq r^2$ . Hence,  $k \in \mathcal{O}(\log r) \subseteq \mathcal{O}(n)$ . Thus, we may assume  $N \geq 2$  and  $1/N \leq 1/2$ . Since  $e^x < 1 + 2x$  for  $0 < x \leq 1$ , it is enough to estimate an upper bound of  $k$  using the bound  $e^{k/N} \leq r^2$ . This implies  $k \in \mathcal{O}(Nn)$ . Thus, in  $\text{NTIME}(\mathcal{O}(Nn))$  we can guess a sequence of length  $k$  of the form

$$r \in g_1 L(\mathcal{A}_1) g_2 L(\mathcal{A}_2) \cdots g_k L(\mathcal{A}_k). \quad (13)$$

Here,  $r$  is as in (13), the  $g_j$ -th belong to  $T$ , and each  $\mathcal{A}_i$  is a subautomaton of some  $\mathcal{A}_j$  mentioned in (13). As a consequence, the NFA  $\mathcal{B}$  is flat over  $\text{GL}(2, \mathbb{Z})$  and it satisfies

$$r \in g_1 L(\mathcal{A}_1) g_2 L(\mathcal{A}_2) \cdots g_k L(\mathcal{A}_k) \iff r \in L(\mathcal{A}_B). \quad (14)$$

Theorem 35 implies that we can decide whether  $r \in L(\mathcal{A}_B)$  in doubly deterministic exponential time  $\text{DTIME}(2^{\mathcal{O}(Nn)})$ . We have  $\text{NTIME}(\mathcal{O}(Nn)) \subseteq \text{DTIME}(2^{\mathcal{O}(Nn)})$ . However, running two  $\text{DTIME}(2^{\mathcal{O}(Nn)})$  one after another is still in  $\text{DTIME}(2^{\mathcal{O}(Nn)})$ . In worst case, we have  $N \in 2^{\mathcal{O}(n)}$  and  $n$  was a polynomial in the original input size. The result follows.  $\blacktriangleleft$

## 7 Nonsingular and singular matrices

Recall that  $s_q$  for  $q \in \mathbb{Z}$  denotes the matrix  $\begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}$ . The most general problem in the paper where we have positive decidability results is the next one.

▷ **Problem 4.** INPUT: A matrix  $g \in \mathbb{Q}^{2 \times 2}$  and an  $\mathbb{Q}^{2 \times 2}$ -NFA  $\mathcal{A}$  which is flat over the set

$$\text{GL}(2, \mathbb{Z}) \cup \{r \in \mathbb{Q} \mid |r| > 1\} \cup \{s_0\} \text{ if } g \neq 0 \quad (15)$$

$$\text{GL}(2, \mathbb{Z}) \cup \mathbb{Q} \cup \{s_0\} \text{ if } g = 0 \quad (16)$$

QUESTION: “ $g \in L(\mathcal{A})$ ?”

Concerning Problem 4 the aim of this section is to prove the following theorem.

► **Theorem 37.** *Let  $\|g\|_{\text{bin}} + \|\mathcal{A}\|_{\text{bin}}$  be the (binary) input size in Problem 4. Then there is an NP-reduction of Problem 4 to polynomially many instances of Problem 1. In particular, Problem 4 is decidable in deterministic doubly exponential time  $\text{DTIME}(2^{2^p})$  where  $p$  is some polynomial.*

### 7.1 Proof of Theorem 37

The following lemma is used below in Section 7.1.1 when we reduce (20) to (24).

► **Lemma 38.** *Let  $m \in \mathbb{Z}$  and  $\begin{pmatrix} a & a' \\ b & b' \end{pmatrix}, \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} \in \mathbb{Z}^{2 \times 2}$  such that*

$$\begin{pmatrix} m & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & a' \\ b & b' \end{pmatrix} s_0 \begin{pmatrix} c & d \\ c' & d' \end{pmatrix}. \quad (17)$$

*Then  $ac = m$ . For  $m \neq 0$  we have  $\begin{pmatrix} a & a' \\ b & b' \end{pmatrix} s_0 = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$  and  $s_0 \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & 0 \end{pmatrix}$ . For  $m = 0$  we have  $\begin{pmatrix} a & a' \\ b & b' \end{pmatrix} s_0 = 0$  or  $s_0 \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} = 0$ .*

**Proof.** We have

$$\begin{pmatrix} a & a' \\ b & b' \end{pmatrix} s_0 \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} = \begin{pmatrix} a & a' \\ b & b' \end{pmatrix} s_0 s_0 \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} = \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}$$

We conclude  $ac = m$  and  $ad = bc = bd = 0$ . For  $m \neq 0$  this implies  $a \neq 0 \neq c$  and therefore  $b = d = 0$ . For  $m = 0$  this implies  $a = 0$  or  $c = 0$ . Say, by symmetry,  $a = 0$ . If  $b = 0$ , we are done. Hence, we may assume  $b \neq 0$ . This implies  $d = 0$ . The lemma follows. ◀

► **Definition 39.** *For  $a \in \mathbb{Z}$  we let  $M_{ij}(a) = \{ \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \in \text{GL}(2, \mathbb{Z}) \mid g_{ij} = a \}$ .*

▷ **Problem 5.**

INPUT: An integer  $a \in \mathbb{Z}$  and a  $\text{GL}(2, \mathbb{Q})$ -NFA  $\mathcal{A}$  which is flat over  $\text{GL}(2, \mathbb{Z})$ : The integer  $a$  and matrix entries are written in binary.

QUESTION: “ $M_{11}(a) \cap L(\mathcal{A}) \neq \emptyset$ ?”

► **Lemma 40.** *There is a NP reduction of Problem 5 to Problem 1. In particular, we can solve Problem 5 in  $\text{DEXPTIME}$  by Theorem 35.*

**Proof.** W.l.o.g.  $\mathcal{A}$  is a  $(\text{GL}(2, \mathbb{Q}) \cap \mathbb{Z}^{2 \times 2})$ -NFA with input size  $n = \|a\|_{\text{bin}} + \|\mathcal{A}\|_{\text{bin}}$ . We have  $M_{11}(a) \cap L(\mathcal{A}) \neq \emptyset$  if and only if there  $b, c, d \in \mathbb{Z}$  such that  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in L(\mathcal{A})$ . Since  $\mathcal{A}$  is flat over  $\text{GL}(2, \mathbb{Z})$  the determinant  $D = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  is bounded by some polynomial in  $n$ . Hence, we can guess  $D$ . For  $a = 0$  we have  $bc = D$  and we can guess  $b$  and  $c$  and then compute  $0 \leq d' \leq |c|$  such that

$$\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \in L(\mathcal{A}) \iff \begin{pmatrix} 0 & b \\ c & d' \end{pmatrix} \in L(\mathcal{A}) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\mathbb{Z}}.$$

The question “ $\begin{pmatrix} 0 & b \\ c & d' \end{pmatrix} \in L(\mathcal{A}) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\mathbb{Z}}$ ?” is an instance of Problem 1. Thus, we may assume  $a \neq 0$ . As a consequence, there are integers  $b', c', d'$  with  $0 \leq |b'|, |c'| \leq |a|$  such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in L(\mathcal{A}) \iff \begin{pmatrix} a & b' \\ c' & d' \end{pmatrix} \in \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{\mathbb{Z}} L(\mathcal{A}) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\mathbb{Z}}.$$

Since  $d' = (D + b'c')/a$ , the binary size of  $d'$  is polynomially bounded in  $n$ . Thus, we can guess the integers  $b', c'$  within NP and compute  $d'$ . We conclude as we did for  $a = 0$ , and we are done. ◀

### 7.1.1 Preprocessing

For a nonsingular input matrix  $g$  we are done by Theorem 36. Indeed, a singular matrix cannot appear on any accepting path and  $\{r \in \mathbb{Q} \mid |r| > 1\} \subseteq \{g \in \text{GL}(2, \mathbb{Q}) \mid |\det(g)| > 1\}$ . Thus, henceforth in the proof we assume that the input matrix  $g$  is singular.

Throughout, we denote by  $S_{\text{sing}}$  and  $S_{0,\text{sing}}$  the following subsets of  $\mathbb{Q}^{2 \times 2}$ :

$$S_{\text{sing}} = \text{GL}(2, \mathbb{Z}) \cup \{r \in \mathbb{Q} \mid |r| > 1\} \cup \{s_0\} \quad (18)$$

$$S_{0,\text{sing}} = \text{GL}(2, \mathbb{Z}) \cup \{s_0\} \quad (19)$$

We begin with a preprocessing.

1. In a first step modify the input such that all matrices are in  $\mathbb{Z}^{2 \times 2}$ .
2. In a second step we compute in deterministic polynomial time for every label of a transition its Smith normal form and split the corresponding transitions into three. In particular,, we can write  $g = e_0 \begin{pmatrix} m & 0 \\ 0 & 0 \end{pmatrix} f_0$  where  $m \in \mathbb{N}$  and  $e_0 f_0 \in \text{GL}(2, \mathbb{Z})$ .
3. Making guesses within **NP** we assume without restriction that we start with the question whether the following assertion is true.

$$e_0 \begin{pmatrix} m & 0 \\ 0 & 0 \end{pmatrix} f_0 \in L(\mathcal{A}_1) e_1 \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} f_1 \cdots L(\mathcal{A}_t) e_t \begin{pmatrix} 1 & 0 \\ 0 & q_t \end{pmatrix} f_t. \quad (20)$$

Here,  $t \geq 1$  and for all  $i$  we have  $e_i, f_i \in \text{GL}(2, \mathbb{Z})$ , and  $q_i \in \mathbb{Z}$ . Moreover, we can assume that following additional properties.

- Each  $\mathcal{A}_i$  is a trim  $S_{\text{sing}}$ -NFA.
  - The assertion (20) holds. If not, then we will make sure that every nondeterministic run of the algorithm will either make no output at all or it will say “No”.
  - The NFA  $\mathcal{A}$  is without any transition labeled by 0. In particular,  $\mathcal{A}$  is flat over  $S_{\text{sing}}$ . This is clear for  $m \neq 0$ . For  $m = 0$  we do the contrary: if some transition labeled by 0 appears, then we accept immediately because  $\mathcal{A}$  is trim.
  - Whenever the label  $a'$  of a transition  $p \xrightarrow{s} q$  is an integer, then  $a'$  divides  $m$ . For  $m = 0$  we make an additional modification. Whenever the label of a transition  $p \xrightarrow{r} q$  is a rational number, then we replace the label by  $1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Thus,  $\mathcal{A}$  is a  $S_{0,\text{sing}}$ -NFA for  $g = 0$ , and we keep this as an invariant.
4. In deterministic polynomial time (and with a renaming) we reduce (20) to the following problem.

$$\begin{pmatrix} m & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{A}_1) \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} \cdots L(\mathcal{A}_t) \begin{pmatrix} 1 & 0 \\ 0 & q_t \end{pmatrix}. \quad (21)$$

The notation and the conditions are as for (20). Moreover, we may assume  $|q_i| \geq 2$  for all  $1 \leq i \leq t$ . Note that  $\begin{pmatrix} m & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{A}_1) \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} \cdots L(\mathcal{A}_t)$  implies that that at least one matrix  $s_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  is used on every accepting path. If  $m = 0$  then every accepting path must use it at least twice.

5. Nondeterministically we guess a transition inside some  $\mathcal{A}_i$  which is labeled by  $s_0$ ; and we guess a factorization  $m = a' a''$  with  $a', a'' \in \mathbb{Z}$ . (For  $m = 0$  it is enough to guess whether  $a' = 0$  or  $a'' = 0$ .) Next, we apply Lemma 38. Since  $\mathcal{A}$  is trim and  $s_0^2 = s_0$ , we can reduce (21) for  $m \neq 0$  to the following two problems (and for  $m = 0$  to the first problem if  $a' = 0$  and otherwise to the second problem):

$$\begin{pmatrix} a' & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{A}'_1) \begin{pmatrix} 1 & 0 \\ 0 & q'_1 \end{pmatrix} \cdots L(\mathcal{A}'_t) \begin{pmatrix} 1 & 0 \\ 0 & q'_t \end{pmatrix} s_0 \quad (22)$$

$$\begin{pmatrix} a'' & 0 \\ 0 & 0 \end{pmatrix} \in s_0 L(\mathcal{A}''_1) \begin{pmatrix} 1 & 0 \\ 0 & q''_1 \end{pmatrix} \cdots L(\mathcal{A}''_t) \begin{pmatrix} 1 & 0 \\ 0 & q''_t \end{pmatrix} \quad (23)$$

By left-right symmetry, it is enough to consider the second problem in (23). After a renaming we have reduced (20) to the following problem

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in s_0 L(\mathcal{A}_1) \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} \cdots L(\mathcal{A}_t) \begin{pmatrix} 1 & 0 \\ 0 & q_t \end{pmatrix}. \quad (24)$$



The notation and conditions are as in (21). Note that in (24) actually every accepting path start with a transition labeled by  $s_0$ . By a abuse of language we denote the NFA corresponding to the right-hand side in (24) by  $\mathcal{A}$  again. Thus, in the new notation we begin with the assertion  $a s_0 \in L(\mathcal{A})$  with  $a \in \mathbb{Z}$ ; and we have to verify it nondeterministically.

This finishes the preprocessing phase.

► **Lemma 41.** *The preprocessing can be realized by an NP-reduction. That is, in order to prove Theorem 37 it is enough to assume that the question in Problem 4 is given in the form*

$$\text{“} \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in s_0 L(\mathcal{A}_1) \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} \cdots L(\mathcal{A}_t) \text{?”} \quad (25)$$

where according to (24) we have  $a \in \mathbb{N}$ .

**Proof.** Clearly, if  $a < 0$ , then we can modify  $\mathcal{A}_1$  to replace the  $a$  by  $-a \in \mathbb{N}$ . Expecting the preprocessing it is clear that time complexity of every run of the nondeterministic procedure is bounded by some polynomial. Moreover, if the answer to the question in Problem 4 is “Yes”, then there is at least one output with a “Yes”-answer to the question in (25). Finally, if the answer to the question in Problem 4 is “No”, then there no nondeterministic output with a “Yes”-answer to that question. ◀

### 7.1.2 After preprocessing

Lemma 41 tells us that it is enough to prove Theorem 37 in the special case where the question is “ $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$ ?” with  $a \in \mathbb{N}$  and  $\mathcal{B}$  denotes the NFA corresponding to the flat rational expression in Equation (24)

$$s_0 L(\mathcal{A}_1) \begin{pmatrix} 1 & 0 \\ 0 & q_1 \end{pmatrix} \cdots L(\mathcal{A}_t) \begin{pmatrix} 1 & 0 \\ 0 & q_t \end{pmatrix}.$$

Without restriction the following conditions are satisfied.

1. The NFA  $\mathcal{B}$  is trim.
2. We have  $t \geq 1$  and for each  $1 \leq i \leq t$  every label of every transition in  $\mathcal{A}_i$  is in  $S_{\text{sing}}$  for  $a \neq 0$  and in the (smaller set)  $S_{0,\text{sing}}$  for  $a = 0$ . In particular, there is no transition with label 0.
3. We have  $|q_i| \geq 2$  for  $1 \leq i < t$  and  $q_t = 0$  for  $a = 0$  and  $q_t = 1$ , otherwise.
4. There is a unique initial state  $p_0$  and a unique final state  $p'_t$  such that  $p_0 \neq p'_t$ .
5. The initial state  $p_0$  has no incoming transitions and all outgoing transitions are labeled by  $s_0$ .
6. The final state  $p'_t$  has no outgoing transition and exactly one incoming transition  $p_t \xrightarrow{g_t} p'_t$  where  $g_t = \begin{pmatrix} 1 & 0 \\ 0 & q_t \end{pmatrix}$ .

For states  $p, q$  in  $\mathcal{B}$ , we denote by  $\mathcal{B}[p, q]$  the maximal subautomaton of  $\mathcal{B}$  which defined by the following conditions

- There is a unique minimal state  $p$  and a unique final state  $q$ .
- Every label of a transition in  $\mathcal{B}$  belongs to  $\text{GL}(2, \mathbb{Q})$ . That is, all transitions with label  $s_0$  are removed, and therefore  $L(\mathcal{B}[p, q]) \subseteq \text{GL}(2, \mathbb{Z})$ .

We are interested only in the automata  $\mathcal{B}[p_1, q]$  where  $L(\mathcal{B}[p, q]) \neq \emptyset$  and  $q \neq p'_t$ .

We now perform a Benois-type of flooding the NFA  $\mathcal{B}$  by adding polynomially many  $s_0$ -transitions. Formally we run the following procedure which transforms the pair  $(a, \mathcal{B})$ .

**procedure** FLOODING( $a, \mathcal{B}$ )

Repeat the following loop until the pair  $(a, \mathcal{B})$  stabilizes

**loop**

- If necessary, make  $\mathcal{B}$  smaller such that  $\mathcal{B}$  becomes trim.

- Guess that  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$ .
- Let  $\pi$  denote a path from a state to the state  $p_t$  such that the path  $(p_0 \xrightarrow{s_0} p_1, \pi, p_t \xrightarrow{g_t} p'_t)$  accepts  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ . Moreover, among all possible choices, let  $\pi$  minimize the number of transitions on  $\pi$  which are labeled by  $s_0$ . Guess, whether the number is zero. If the guess is zero, then **exit the loop** with the current pair  $(a, \mathcal{B})$ .  
Thus, we can assume that  $\pi$  uses a transition  $p \xrightarrow{s_0} q$ . We guess that transition; and we write  $\pi = (\pi', p \xrightarrow{s_0} q, \pi'')$ .
- Guess the state  $p$  and verify  $L(\mathcal{B}[p_1, p]) \neq \emptyset$ . Note that such a state  $p$  exists because we can guess the first transition on  $\pi$  which is labeled by  $s_0$ .
- Suppose in this item that  $a = 0$ .  
Check whether  $M_{11}(0) \cap L(\mathcal{B}[p_1, p]) \neq \emptyset$ . If the answer is “yes”, then we stop the procedure with the answer: “Yes, we have  $0 \in L(\mathcal{B})$ ”.  
If the answer is “no”, then there exists some  $0 \neq a \in \mathbb{Z}$  such that  $M_{11}(a) \cap L(\mathcal{B}[p_1, p]) \neq \emptyset$  and we introduce an  $s_0$ -transition  $p_0 \xrightarrow{s_0} q$ . This is a new transition because the number of transitions on  $\pi$  which are labeled by  $s_0$  is minimal, and  $(p_0 \xrightarrow{s_0} q, \pi'', p_t \xrightarrow{g_t} p'_t)$  is an accepting path for 0 using less  $s_0$ 's than  $\pi$  since the additional  $s_0$ -transition allows a shortcut for accepting 0.  
If the procedure did not stop, then let  $\mathcal{B}'$  be the NFA with one more  $s_0$ -transition than  $\mathcal{B}$ . By construction we have  $0 \in L(\mathcal{B}) \iff 0 \in L(\mathcal{B}')$ . Rename  $\mathcal{B}'$  as  $\mathcal{B}$ .
- Suppose in this item that  $a \neq 0$ .  
Since  $a > 0$  there must be some integer matrix  $\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in L(\mathcal{B}[p_1, p])$  where  $0 < a'$  divides  $a$  and  $\| \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \|_{\text{bin}}$  can be guessed in **NP** by Lemma 40.  
For a moment introduce a new transition  $p_0 \xrightarrow{a's_0} q$ . For  $a' = 1$  it is a new transition because (just as in the case above where  $a = 0$ ) the path  $(p_0 \xrightarrow{s_0} q, \pi'', p_t \xrightarrow{g_t} p'_t)$  is accepting for  $as_0$  and  $\pi''$  is using less  $s_0$ 's than  $\pi$ .  
If  $a' \neq 1$ , then let  $a'' = a/a'$ . Note that  $0 < a'' < a$ . Remove all outgoing transitions at the initial state  $p_0$  and after that introduce  $p_0 \xrightarrow{s_0} q$  as the unique outgoing transition at the initial state. Then the path  $(p_0 \xrightarrow{1} q, \pi'', p_t \xrightarrow{g_t} p'_t)$  is an accepting path for  $a''s_0$  with  $|a''| \leq |a|/2$ . Let  $\mathcal{B}'$  be the NFA after these modification. Note that we have the following property.

$$\begin{pmatrix} a'' & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B}') \subseteq L(\mathcal{B}) \quad (26)$$

Rename the pair  $(a'', \mathcal{B}')$  as  $(a, \mathcal{B})$ .

#### endloop

- For  $a = 0$  we check whether there is a transition  $p_0 \xrightarrow{s_0} p_1$  such that  $M_{11}(0) \cap L(\mathcal{B}[p_1, p_t]) \neq \emptyset$ . We can do so by Lemma 40. In the positive case we output “Yes”, otherwise we stop without any output. Indeed, we have  $q_t = 0$  and  $0 = s_0 \begin{pmatrix} a & b \\ c & d \end{pmatrix} s_0 \iff \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{11}(0)$ .
- For  $0 < a$  we ask whether there exists a matrix  $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \in L(\mathcal{B}[p_1, p_t])$  because  $s_0 \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$  and  $q_t = 1$ . Since  $a \in \mathbb{N}$  and  $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \in \text{GL}(2, \mathbb{Z})$ , we conclude  $a = d = 1$ . Note that we have  $\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \in L(\mathcal{B}[p_1, p_t])$  if and only if  $1 \in \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\mathbb{Z}} L(\mathcal{B}[p_1, p_t])$ . we check whether there is a transition  $p_0 \xrightarrow{s_0} p_1$  such that  $M_{11}(0) \cap L(\mathcal{B}[p_1, p_t]) \neq \emptyset$ . We can do so by Lemma 40.  
Thus, for  $0 < a$  we check first that  $a = 1$ . If  $a \neq 1$ , then we stop without any output. If  $a = 1$ , then we check  $1 \in \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\mathbb{Z}} L(\mathcal{B}[p_1, p_t])$ . This is an instance of Problem 1.  
The algorithm returns “Yes” if and only if there is some transition  $p_0 \xrightarrow{s_0} p_1$  such that  $1 \in \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\mathbb{Z}} L(\mathcal{B}[p_1, p_t])$  is true. Otherwise it stops without any output.

#### endprocedure

► **Lemma 42.** Let  $n = \|a\|_{\text{bin}} + \|\mathcal{B}\|_{\text{bin}}$  and  $t(n)$  be an upper bound for worst-case running time of the procedure  $\text{FLOODING}(a, \mathcal{B})$  under the assumption that instances to Problem 1 are answered by some oracle in constant time. Then the following assertions hold.

1. We can bound  $t(n)$  by some polynomial in  $n$ . In particular, every run terminates.

2. If  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$ , then there is at least one successful run answering “Yes”.
3. If  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \notin L(\mathcal{B})$ , then no run answers “Yes”: either no output is produced or the answer is “No”.

**Proof.** Let  $s$  be the number of states of  $\mathcal{B}$ . The procedure never increases  $s$ . In order to find a polynomial upper bound on  $t(n)$  we observe that in each loop either a new  $s_0$ -transition is introduced or  $0 < a$  and we decrease the number  $a$  by at  $a/2$ .

If  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$ , then we are able to guess at each step correctly; and there is at least one run answering “Yes”.

If  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \notin L(\mathcal{B})$ , then, of course, the first guess  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$  is wrong. The procedure forces us to leave the loop where  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$  is still wrong.

Inspecting the behavior of the procedure after the line **endloop** an output “Yes” proves  $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in L(\mathcal{B})$ . Contradiction. ◀

Lemma 42 finishes the proof of Theorem 35 and therefore Section 7.1.

## 8 Conclusion

Decidability of membership in group theory has a long history going back to the work of Dehn (and others) at the beginning of the 20th century. Of particular interest are the membership problems for  $GL(n, \mathbb{Z})$  and  $GL(n, \mathbb{Q})$  but as soon as  $n \geq 3$  various natural decision problems become undecidable, whereas the corresponding problems remain open for  $n = 2$ .

The contributions of the paper are as follows. On a conceptual level, we introduce the notion of a family of flat rational sets  $\text{Frat}(M, N)$  with respect to a semigroup  $M$  and a submonoid  $N$ . It turns out that  $\text{Frat}(M, N)$  contains  $\text{Rat}(N)$ , and it is a subfamily of  $\text{Rat}(M)$ . For us, the most interesting case is when  $N = H$  is group. In this case  $\text{Frat}(M, N)$  has an inductive definition which is without reference to a particular presentation of  $M$  or  $H$ , see Theorem 22.

Another main contribution is the dichotomy stated in Theorem 17. It shows that if a subgroup  $G$  of  $GL(2, \mathbb{Q})$  contains  $GL(2, \mathbb{Z})$  and a diagonal but not central matrix  $g$ , then  $G$  contains a Baumslag-Solitar group  $BS(1, q)$  where  $q \geq 2$  and the Baumslag-Solitar group has infinite index in  $G$ . As a consequence there is no hyperbolic subgroup in  $GL(2, \mathbb{Q})$  which has  $GL(2, \mathbb{Z})$  as a proper subgroup. In particular, with respect to inclusion,  $GL(2, \mathbb{Z})$  is a maximal virtually free group in  $GL(2, \mathbb{Q})$ .

There is a natural hierarchy of decision problems.

1. Membership to f.g. subgroups.
2. Membership to f.g. submonoids.
3. Membership to rational subsets.
4. Equality of rational subsets.

For  $GL(2, \mathbb{Z})$ , equality of rational subsets is decidable. The dichotomy implies that for any subgroup  $G$ , which is larger than  $GL(2, \mathbb{Z})$ , either membership to rational subsets is decidable but equality of rational subsets is undecidable or, in the other case, we don’t know (when the paper is written) whether membership to f.g. subgroups for  $G$  is decidable. These facts were main reasons to define the notion of a flat rational set. It pushes the positive decidability results for  $GL(2, \mathbb{Z})$  further to the relative Boolean algebra  $\text{Frat}(GL(2, \mathbb{Q}), GL(2, \mathbb{Z}))$  (and beyond if we include nonsingular matrices). Using several structural results of flat rational sets, the main positive decidability result Theorem 37 being a generalization of Theorem 36 since it includes singular matrices.

Lines for future research include the following topics. (1) Find other natural applications using membership problems for flat rational sets. For example, when considering  $GL(2, k)$  where  $k$  is an algebraic field over  $\mathbb{Q}$  or a function field in one variable over a finite field. (2) Is the membership problem for f.g. subgroups in any subgroup  $G$  of  $GL(2, \mathbb{Q})$  decidable if  $GL(2, \mathbb{Z}) \leq G$  and  $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \in G$  where  $p$  is prime? (3) Several statements of our paper contain complexity bounds but we don’t know whether they are sharp. For example, we don’t know whether Problem 1 is **NP**-complete.

## References

- 1 A. V. Anisimow and F. D. Seifert. Zur algebraischen Charakteristik der durch kontext-freie Sprachen definierten Gruppen. *Elektron. Informationsverarbeitung. Kybernetik*, 11(10–12):695–702, 1975.
- 2 L. Babai, R. Beals, J.-Y. Cai, G. Ivanyos, and E. M. Luks. Multiplicative equations over commuting matrices. In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '96*, pages 498–507, Philadelphia, PA, USA, 1996. Society for Industrial and Applied Mathematics.
- 3 G. Baumslag and D. Solitar. Some two-generator one-relator non-Hopfian groups. *Bull. Amer. Math. Soc.*, 68:199–201, 1962.
- 4 H. Behr and J. Mennicke. A presentation of the groups  $\text{PSL}(2, p)$ . *Canadian Journal of Mathematics*, 20:1432–1438, 1968.
- 5 P. Bell, V. Halava, T. Harju, J. Karhumäki, and I. Potapov. Matrix equations and Hilbert’s tenth problem. *International Journal of Algebra and Computation*, 18:1231–1241, 2008.
- 6 P. Bell, I. Potapov, and P. Semukhin. On the mortality problem: From multiplicative matrix equations to linear recurrence sequences and beyond. In *Proc. 44th MFCS, LIPIcs*, pages 83:1–83:15, 2019.
- 7 P. C. Bell, M. Hirvensalo, and I. Potapov. The identity problem for matrix semigroups in  $\text{sl}_2(F)$  is  $\text{np}$ -complete. In P. N. Klein, editor, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 187–206. SIAM, 2017.
- 8 M. Benoist. Parties rationelles du groupe libre. *C. R. Acad. Sci. Paris, Sér. A*, 269:1188–1190, 1969.
- 9 M. Cadilhac, D. Chistikov, and G. Zetsche. Rational subsets of Baumslag-Solitar groups. In A. Czumaj, A. Dawar, and E. Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 116:1–116:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.
- 10 J. Cassaigne, V. Halava, T. Harju, and F. Nicolas. Tighter undecidability bounds for matrix mortality, zero-in-the-corner problems, and more. *arXiv eprints*, abs/1404.0644, 2014.
- 11 É. Charlier and J. Honkala. The freeness problem over matrix semigroups and bounded languages. *Inf. Comp.*, 237:243–256, 2014.
- 12 T. Colcombet, J. Ouaknine, P. Semukhin, and J. Worrell. On reachability problems for low-dimensional matrix semigroups. In C. Baier, I. Chatzigiannakis, P. Flocchini, and S. Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, pages 44:1–44:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
- 13 V. Diekert and M. Elder. Solutions to twisted word equations and equations in virtually free groups. *International Journal of Algebra and Computation*, 30:731–819, 2020. Based on the conference in *LIPIcs.ICALP.2017.96:1–96:14*.
- 14 V. Diekert, M. Kufleitner, G. Rosenberger, and U. Hertrampf. *Discrete Algebraic Methods. Arithmetic, Cryptography, Automata and Groups*. Walter de Gruyter, 2016.
- 15 V. Diekert and G. Rozenberg, editors. *The Book of Traces*. World Scientific, Singapore, 1995.
- 16 C. Druţu and M. Kapovich. *Geometric Group Theory*, volume 63 of *Colloquium Publications*. American Mathematical Society, Providence (RI), 2018.
- 17 S. Eilenberg. *Automata, Languages, and Machines*, volume A. Academic Press, New York and London, 1974.
- 18 S. Eilenberg and M.-P. Schützenberger. Rational sets in commutative monoids. *Journal of Algebra*, 13:173–191, 1969.
- 19 S. M. Gersten. Dehn functions and  $l_1$ -norms of finite presentations. In *Algorithms and classification in combinatorial group theory (Berkeley, CA, 1989)*, volume 23 of *Math. Sci. Res. Inst. Publ.*, pages 195–224. Springer, New York, 1992.
- 20 Z. Grunschlag. *Algorithms in Geometric Group Theory*. PhD thesis, University of California, 1999. AAI9931252.
- 21 Y. Gurevich and P. Schupp. Membership problem for the modular group. *SIAM J. Comput.*, 37:425–459, 2007.
- 22 T. Harju. Post correspondence problem and small dimensional matrices. In *Proc. 13th DLT*, volume 5583 of *LN in Comp. Sci.*, pages 39–46, 2009.

- 23 IJstrand Jan Aalbersberg and H. J. Hoogeboom. Characterizations of the decidability of some problems for regular trace languages. *Mathematical Systems Theory*, 22:1–19, 1989.
- 24 R. Kannan and A. Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. *SIAM*, 8:499–507, 1979.
- 25 S. C. Kleene. Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, number 34 in Annals of Mathematics Studies, pages 3–40. Princeton University Press, 1956.
- 26 S. Ko, R. Niskanen, and I. Potapov. On the identity problem for the special linear group and the heisenberg group. In I. Chatzigiannakis, C. Kaklamanis, D. Marx, and D. Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPICs*, pages 132:1–132:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.
- 27 D. König, M. Lohrey, and G. Zetsche. Knapsack and subset sum problems in nilpotent, polycyclic, and co-context-free groups. *arXiv eprints*, abs/1507.05145, 2015.
- 28 M. Lohrey. Subgroup membership in  $gl(2, F)$ . In M. Bläser and B. Monmege, editors, *38th International Symposium on Theoretical Aspects of Computer Science, STACS 2021, March 16-19, 2021, Saarbrücken, Germany (Virtual Conference)*, volume 187 of *LIPICs*, pages 51:1–51:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.
- 29 M. Lohrey and G. Sénizergues. Theories of HNN-extensions and amalgamated products. In M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, editors, *ICALP*, volume 4052 of *Lecture Notes in Computer Science*, pages 504–515. Springer, 2006.
- 30 M. Lohrey and B. Steinberg. The submonoid and rational subset membership problems for graph groups. *Journal of Algebra*, 320(2):728–755, 2008.
- 31 R. Lyndon and P. Schupp. *Combinatorial Group Theory*. Classics in Mathematics. Springer, 2001. First edition 1977.
- 32 A. A. Markov. On certain insoluble problems concerning matrices. *Dokl. Akad. Nauk SSSR*, 57:539–542, 1947.
- 33 J. D. McKnight. Kleene quotient theorem. *Pacific Journal of Mathematics*, pages 1343–1352, 1964.
- 34 K. A. Mihailova. The occurrence problem for direct products of groups. *Dokl. Akad. Nauk SSSR*, 119:1103–1105, 1958. English translation in: *Math. USSR Sbornik*, 70: 241–251, 1966.
- 35 M. Newman. The structure of some subgroups of the modular group. *Illinois J. Math.*, 6:480–487, 1962.
- 36 Ch. H. Papadimitriou. *Computational Complexity*. Addison Wesley, 1994.
- 37 I. Potapov. Reachability problems in matrix semigroups. *Dagstuhl Reports*, 9:95–98, 2019.
- 38 I. Potapov and P. Semukhin. Decidability of the membership problem for  $2 \times 2$  integer matrices. In *Proc. 28th SODA*, pages 170–186, 2017.
- 39 I. Potapov and P. Semukhin. Membership problem in  $GL(2, \mathbb{Z})$  extended by singular matrices. In *Proc. 42nd MFCS*, pages 44:1–44:13, 2017.
- 40 N. S. Romanovskii. Some algorithmic problems for solvable groups. *Algebra i Logika*, 13:26–34, 121, 1974.
- 41 J. Sakarovitch. The “last” decision problem for rational trace languages. In I. Simon, editor, *Proc. 1st Latin American Symposium on Theoretical Informatics (LATIN’92)*, volume 583 of *Lecture Notes in Computer Science*, pages 460–473, Heidelberg, 1992. Springer-Verlag.
- 42 G. Sénizergues. On the rational subsets of the free group. *Acta Informatica*, 33:281–296, 1996.
- 43 G. Sénizergues. Personal communication, 2019.
- 44 J.-P. Serre. *Trees*. Springer, 1980. French original 1977.
- 45 P. V. Silva. Recognizable subsets of a group: finite extensions and the abelian case. *Bulletin of the EATCS*, 77:195–215, 2002.

## 9 Appendix

### 9.1 Missing proofs for Section 2.2

**Proof of Proposition 2.** W.l.o.g. suppose that  $0 \leq b \leq a$ . We describe an algorithm  $\text{gcd}(a, b)$  that outputs  $x$  and  $y$  such that  $ax + by = \text{gcd}(a, b)$  and  $|x|, |y| \leq \max\{a, b\}$ . If  $b \in \{0, a\}$ , then output  $x = 1$  and  $y = 0$  since the  $\text{gcd}(a, b) = a$  in this case.

Suppose  $0 < b < a$ . Compute  $k$  and  $r$  such that  $a = kb + r$ , where  $0 \leq r < b$ . Note that we also have  $r \leq a/2$ . The computation of  $k$  involves an integer division which is known to have the same complexity as multiplication. For example, by Schönhage-Strassen multiplication can be done in time  $\tilde{O}(n)$ . Since we use soft- $\mathcal{O}$ -notation the reduction from integer division to multiplication is easy using Newton's method.

If  $r = 0$ , then return  $x = 0$  and  $y = 1$  since  $\text{gcd}(a, b) = b$ . If  $r > 0$ , then recursively call  $\text{gcd}(b, r)$ . Let  $u, v$  be the outputs of  $\text{gcd}(b, r)$  such that  $bu + rv = \text{gcd}(b, r) = \text{gcd}(a, b)$  and  $|u|, |v| \leq b$ . Then we have

$$\text{gcd}(a, b) = bu + rv = bu + (a - kb)v = av + b(u - kv).$$

The algorithm  $\text{gcd}(a, b)$  outputs  $x = v$  and  $y = u - kv$ . The number  $y = u - kv$  can be computed in time  $\tilde{O}(n)$  since the absolute values of  $u, v$  and  $k$  are bounded by  $a$ , and hence they are  $n$ -bit numbers. We already have  $|x| = |v| \leq b \leq a$ . To estimate  $|y|$ , observe that

$$|y| \leq \frac{|\text{gcd}(a, b) - av|}{b} \leq \frac{\text{gcd}(a, b)}{b} + \frac{a|v|}{b} \leq \frac{\text{gcd}(a, b)}{b} + a.$$

The assumption  $r > 0$  implies that  $\text{gcd}(a, b) < b$ . Thus  $\text{gcd}(a, b)/b < 1$ . Since  $|y|$  and  $a$  are integers, it follows that  $|y| \leq a$  as required.

It is not hard to see that  $\text{gcd}(a, b)$  requires  $\mathcal{O}(n)$  recursive calls because as we noted above  $r \leq a/2$ . Since each call takes  $\tilde{O}(n)$  time, the total running time is  $\tilde{O}(n^2)$ . ◀

It is possible to improve the running time to compute the values  $x, y$  in time  $\mathcal{O}(n^2)$ . But this is not important since for the following results any polynomial time bound would suffice.

**Proof of Lemma 1.** For  $|q| = 1$ , the numbers  $x = y = 1$  are coprime; and they satisfy  $xb + yd \equiv 0 \pmod{q}$  because all integers are congruent modulo 1. Hence, we may assume  $2 \leq |q|$ .

Let  $P_1$  be the set of primes  $p$  such that  $\text{gcd}(p, d) = 1$  and  $P_2$  the set of all other primes. That is  $p \in P_2$  implies  $p \mid d$ . Write  $q = q_1 \cdot q_2$  such that  $q_i$  uses primes from  $P_i$ , only. For every prime  $p$  we have

$$p \in P_1 \implies \text{gcd}(p, d) = 1 \tag{27}$$

$$p \in P_2 \implies \text{gcd}(p, b) = 1, \text{ because } \text{gcd}(b, d) = 1. \tag{28}$$

Hence,  $d$  is invertible in  $\mathbb{Z}/q_1$  and  $b$  is invertible in  $\mathbb{Z}/q_2$ . Therefore we can solve the following.

$$x_1 \equiv 1 \pmod{q_1} \text{ and } y_1 \equiv -bd^{-1} \pmod{q_1} \tag{29}$$

$$y_2 \equiv 1 \pmod{q_2} \text{ and } x_2 \equiv -db^{-1} \pmod{q_2} \tag{30}$$

Since  $\text{gcd}(q_1, q_2) = 1$  we obtain by Chinese remaindering  $x, y$  with  $1 \leq x, y < |q|$  such that

$$x \equiv 1 \pmod{q_1} \text{ and } x \equiv -db^{-1} \pmod{q_2} \tag{31}$$

$$y \equiv 1 \pmod{q_2} \text{ and } y \equiv -bd^{-1} \pmod{q_1} \tag{32}$$

The congruences in (31) and (32) tell us that these  $x, y$  with  $1 \leq x, y < |q|$  satisfy

$$xb + yd \equiv 0 \pmod{q}. \tag{33}$$

Indeed, the congruence in (33) holds mod  $q_1$  and mod  $q_2$ , hence it holds mod  $q$ .

We claim that  $\gcd(x, y, q) = 1$ . To see this, let  $p \mid q$ . Then  $p \mid q_i$  for exactly one  $i \in \{1, 2\}$ . Say  $i = 1$ , then  $x \equiv 1 \pmod{q_1}$  implies  $x \equiv 1 \pmod{p}$  because  $p \mid q_1$ . Hence,  $\gcd(x, q_1) = 1$ . For  $i = 2$  we obtain  $\gcd(y, q_2) = 1$  and therefore  $\gcd(x, y, q) = 1$  since  $q = q_1 q_2$ . Thus, the claim.

It is still possible that there is a prime  $p$  such that  $p \mid x$  and  $p \mid y$ . However, since  $\gcd(x, y, q) = 1$  such a prime  $p$  is invertible in  $\mathbb{Z}/q$ . Thus,

$$\frac{x}{p}b + \frac{y}{p}d \equiv 0 \pmod{q}. \quad (34)$$

The property  $\gcd(\frac{x}{p}, \frac{y}{p}, q) = 1$  is inherited. So we can make  $x$  and  $y$  smaller. We are done.  $\blacktriangleleft$

## 9.2 Generators of $\mathrm{SL}(2, \mathbb{Z}[1/p])$

**First**, we prove a claim used in the proof of Corollary 12 which states that if  $H$  is a subgroup of finite index in a f.g. group  $G$  such the word problem for  $H$  is decidable, then the word problem for  $G$  is decidable.

Since  $G/H$  is finite,  $H$  is f.g.. This is classical fact, see e. g. [31]. Having this, assume we can decide the word problem for  $H$  in time  $t(n)$  for words on length  $n$  over some finite generating set  $A$  for  $H$ . Let  $R \subseteq G$  be a subset of size  $|G/H| - 1$  such that  $G \setminus H = HR \cup H$ . Then  $A \cup R$  generates  $G$ . Moreover for each  $a \in A \cup R$  and  $r \in R$  we find a word  $v_{a,r} \in A^*$  and an element  $s_{a,r} \in R$  such that  $ra = v_{a,r}s_{a,r}$  in  $G$ . By moving coset representatives to the right we can transform any word  $w$  of length  $n$  over  $A \cup R$  to a word  $w'$  in  $A^* \cup A^*R$ . The length of the word  $w'$  is bounded by  $cn$  where  $c = \max\{|v_{a,r}| \mid a \in A \cup R, r \in R\}$ . For  $w = 1 \in G$  we must have  $w' \in A^*$ . Thus we can check whether  $w' = 1$  in time  $t(cn)$ . Changing to another generating set for  $G$  leads to time  $t(\mathcal{O}(n))$ .

**Finally**, we give a simple proof for the well-known fact that  $\mathrm{SL}(2, \mathbb{Z}[1/p])$  is generated by  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ , and  $\begin{pmatrix} p & 0 \\ 0 & p^{-1} \end{pmatrix}$ . As usual,  $\mathbb{Z}[1/p]$  is the ring  $\{p^n r \in \mathbb{Q} \mid n, r \in \mathbb{Z}\}$ . We use the following notation: let  $\alpha, \beta, \gamma, \delta$  denote elements in  $\mathbb{Z}[1/p]$ , and  $a, b, c, d$  denote elements in  $\mathbb{Z}$ . Starting with a matrix  $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$  we do the following:

1. Multiply by  $\begin{pmatrix} p^{-1} & 0 \\ 0 & p \end{pmatrix}$  on the left until we reach  $\begin{pmatrix} \alpha & \beta \\ c & d \end{pmatrix}$ .
2. Multiply by  $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ ,  $\begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}$ , and  $\begin{pmatrix} 1 & 0 \\ \pm 1 & 1 \end{pmatrix}$  until we reach  $\begin{pmatrix} \alpha & \beta \\ c & d \end{pmatrix}$ . This is trivial for  $|c| = |d|$ . In the other case we may assume  $|c| > |d|$ . Next, transform  $\begin{pmatrix} \alpha & \beta \\ c & d \end{pmatrix}$  into a matrix of type  $\begin{pmatrix} \alpha & \beta \\ c \pm d & d \end{pmatrix}$  such that  $|c \pm d| < |c|$ . Use induction on  $|c| + |d|$ .
3. Multiply by  $\begin{pmatrix} p & 0 \\ 0 & p^{-1} \end{pmatrix}$  on the left until we reach  $\begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix}$ .
4. Now,  $\alpha\delta = 1$ . Hence  $\alpha = p^m a$  and  $\delta = p^n d$  where  $\gcd(a, p) = \gcd(d, p) = 1$ . Since  $p$  is a prime,  $m + n = 0$  and  $ad = 1$ .
5. WLOG  $a = d = 1$  and  $m \geq 1$  and hence,  $\begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} = \begin{pmatrix} p^m & \beta \\ 0 & p^{-m} \end{pmatrix}$ .
6. Using  $\begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}$  we can add or subtract the lower row  $p^m |b|$  times to the upper row. Since  $m \geq 1$  we obtain  $\begin{pmatrix} p & 0 \\ 0 & p^{-1} \end{pmatrix}^m$ .